

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:01:39 / Search time 1014.74 Seconds  
(without alignments)  
932.058 Million cell updates/sec

Title: US-09-578-453-2  
Perfect score: 20  
Sequence: 1 GGACATCCCGGCATGTCC 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 500 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	20	100.0	20 6 AR093067	AR093067 Sequence
2	20	100.0	20 6 AR093067	AR093067 Sequence
3	20	100.0	20 6 E32518	E32518 Scavenger r
4	20	100.0	20 6 E32518	E32518 Scavenger r
5	20	100.0	20 6 AR359569	AR359569 Sequence
6	20	100.0	20 6 AR359569	AR359569 Sequence
7	20	100.0	20 6 AX339212	AX339212 Sequence
8	20	100.0	20 6 AX339212	AX339212 Sequence
9	20	100.0	20 9 HSDMP53	HSDMP53 H.sapiens D
10	20	100.0	20 9 HSDMP53	HSDMP53 H.sapiens D
11	20	100.0	24 6 AX472546	AX472546 Sequence
12	20	100.0	24 6 AX472546	AX472546 Sequence
13	20	100.0	24 6 AX476842	AX476842 Sequence
14	20	100.0	24 6 AX476842	AX476842 Sequence
15	20	100.0	24 6 AX476861	AX476861 Sequence
16	20	100.0	24 6 AX476861	AX476861 Sequence
17	20	100.0	56 6 BD272327	BD272327 Selective
18	20	100.0	56 6 BD272327	BD272327 Selective
19	20	100.0	56 6 BD272328	BD272328 Selective

20	20	100.0	56 6 BD272328	BD272328 Selective
21	22	18.4	92.0 637 11 BV056035	BV056035 S212P6623
22	21	18.4	92.0 637 11 BV056035	BV056035 S212P6623
23	23	18.4	92.0 146698	AC021460 Homo sapi
24	24	18.4	92.0 146698	AC021460 Homo sapi
25	25	18.4	92.0 159478	AL359638 Homo sapi
26	26	18.4	92.0 159478	AL359638 Homo sapi
27	27	18.4	92.0 165090	HS44420 Homo sapi
28	28	18.4	92.0 165090	HS44420 Homo sapi
29	29	18.4	92.0 170132	AC126323 Homo sapi
30	30	18.4	92.0 170132	AC126323 Homo sapi
31	31	18.4	92.0 173987	AL670675 Mouse DNA
32	32	18.4	92.0 173987	AL670675 Mouse DNA
33	33	18.4	92.0 185592	AL773522 Mouse DNA
34	34	18.4	92.0 185592	AL773522 Mouse DNA
35	35	18.4	92.0 191128	AC139327 Mus muscu
36	36	18.4	92.0 191128	AC139327 Mus muscu
37	37	18.4	92.0 202227	AC068126 Homo sapi
38	38	18.4	92.0 202227	AC068126 Homo sapi
39	39	18.4	92.0 203950	AC087483 Homo sapi
40	40	18.4	92.0 203950	AC087483 Homo sapi
41	41	18.4	92.0 218688	AC132452 Mus muscu
42	42	18.4	92.0 218688	AC132452 Mus muscu
43	43	18.4	92.0 260120	AC113311 Mus muscu
44	44	18.4	92.0 260120	AC113311 Mus muscu
45	45	18	90.0 233077 10 AL807771	AL807771 Mouse DNA
46	46	18	90.0 233077 10 AL807771	AL807771 Mouse DNA
47	47	18	90.0 263507 2 AC093466	AC093466 Mus muscu
48	48	18	90.0 263507 2 AC093466	AC093466 Mus muscu
49	49	17.4	87.0 1439 1 AY395415	AY395415 Unculture
50	50	17.4	87.0 1439 1 AY395415	AY395415 Unculture
51	51	17.4	87.0 144301 9 AC010467	AC010467 Homo sapi
52	52	17.4	87.0 144301 9 AC010467	AC010467 Homo sapi
53	53	17.4	87.0 165662 9 AC092070	AC092070 Homo sapi
54	54	17.4	87.0 165662 9 AC092070	AC092070 Homo sapi
55	55	17	85.0 551 11 G88857	G88857 S208P644FB
56	56	17	85.0 551 11 G88857	G88857 S208P644FB
57	57	16.8	84.0 30 6 AR064395	AR064395 Sequence
58	58	16.8	84.0 30 6 AR064395	AR064395 Sequence
59	59	16.8	84.0 30 6 AR123186	AR123186 Sequence
60	60	16.8	84.0 30 6 AR123186	AR123186 Sequence
61	61	16.8	84.0 30 6 AR157705	AR157705 Sequence
62	62	16.8	84.0 30 6 AR157705	AR157705 Sequence
63	63	16.8	84.0 30 6 128715	128715 Sequence 18
64	64	16.8	84.0 30 6 128715	128715 Sequence 18
65	65	16.8	84.0 30 6 189369	189369 Sequence 18
66	66	16.8	84.0 30 6 189369	189369 Sequence 18
67	67	16.8	84.0 30 6 AR218599	AR218599 Sequence
68	68	16.8	84.0 30 6 AR218599	AR218599 Sequence
69	69	16.8	84.0 352 10 MMDNABA	MMDNABA
70	70	16.8	84.0 352 10 MMDNABA	MMDNABA
71	71	16.8	84.0 342 11 G91003	G91003
72	72	16.8	84.0 342 11 G91003	G91003
73	73	16.8	84.0 542 11 BV028618	BV028618
74	74	16.8	84.0 542 11 BV028618	BV028618
75	75	16.8	84.0 552 11 BV035566	BV035566 S212P6010
76	76	16.8	84.0 552 11 BV035566	BV035566 S212P6010
77	77	16.8	84.0 593 11 G97714	G97714 S209P6016FH
78	78	16.8	84.0 593 11 G97714	G97714 S209P6016FH
79	79	16.8	84.0 658 11 BV024662	BV024662 S212P6345
80	80	16.8	84.0 658 11 BV024662	BV024662 S212P6345
81	81	16.8	84.0 728 10 BC055943	BC055943 Mus muscu
82	82	16.8	84.0 728 10 BC055943	BC055943 Mus muscu
83	83	16.8	84.0 771 11 BV032844	BV032844 S212P6021
84	84	16.8	84.0 771 11 BV032844	BV032844 S212P6021
85	85	16.8	84.0 816 11 BV071163	BV071163 S212P6013
86	86	16.8	84.0 816 11 BV071163	BV071163 S212P6013
87	87	16.8	84.0 825 11 BV067922	BV067922 S212P6600
88	88	16.8	84.0 825 11 BV067922	BV067922 S212P6600
89	89	16.8	84.0 1565 10 BC002257	BC002257 Mus muscu
90	90	16.8	84.0 1565 10 BC002257	BC002257 Mus muscu
91	91	16.8	84.0 3192 10 MMJ304823	MMJ304823 Mus muscu
92	92	16.8	84.0 3192 10 MMJ304823	MMJ304823 Mus muscu

93	16.8	84.0	3328	10	BC005512	BC005512 Mus muscu	c 166	16.8	84.0	79925	9	AC093719	AC093719 Homo sapi
94	16.8	84.0	3328	10	BC005512	BC005512 Mus muscu	c 167	16.8	84.0	82561	2	AC139202	AC139202 Mus muscu
95	16.8	84.0	9391	6	AR409087	AR409087 Sequence	c 168	16.8	84.0	82561	2	AC139202	AC139202 Mus muscu
96	16.8	84.0	9391	6	AR409087	AR409087 Sequence	c 169	16.8	84.0	82646	10	AL928620	AL928620 Mouse DNA
97	16.8	84.0	9391	6	AX045103	AX045103 Sequence	c 170	16.8	84.0	82646	10	AL928620	AL928620 Mouse DNA
98	16.8	84.0	9391	6	AX045103	AX045103 Sequence	c 171	16.8	84.0	83137	10	AL844573	AL844573 Mouse DNA
99	16.8	84.0	9511	6	AR409086	AR409086 Sequence	c 172	16.8	84.0	83137	10	AL844573	AL844573 Mouse DNA
100	16.8	84.0	9511	6	AR409086	AR409086 Sequence	c 173	16.8	84.0	88526	10	AL807759	AL807759 Mouse DNA
101	16.8	84.0	9511	6	AX045101	AX045101 Sequence	c 174	16.8	84.0	88526	10	AL807759	AL807759 Mouse DNA
102	16.8	84.0	9511	6	AX045101	AX045101 Sequence	c 175	16.8	84.0	90129	10	AL627078	AL627078 Mouse DNA
103	16.8	84.0	9511	10	MMU12147	MMU12147 Mus musculu	c 176	16.8	84.0	90129	10	AL627078	AL627078 Mouse DNA
104	16.8	84.0	9511	10	MMU12147	MMU12147 Mus musculu	c 177	16.8	84.0	93553	10	AL646047	AL646047 Mouse DNA
105	16.8	84.0	15167	2	AC007412	AC007412 Mus muscu	c 178	16.8	84.0	93553	10	AL646047	AL646047 Mouse DNA
106	16.8	84.0	15167	2	AC007412	AC007412 Mus muscu	c 179	16.8	84.0	96727	10	AL627426	AL627426 Mouse DNA
107	16.8	84.0	21565	6	AX695446	AX695446 Sequence	c 180	16.8	84.0	96727	10	AL627426	AL627426 Mouse DNA
108	16.8	84.0	21565	6	AX695446	AX695446 Sequence	c 181	16.8	84.0	96916	10	AL670100	AL670100 Mouse DNA
109	16.8	84.0	22700	2	AC020792	AC020792 Mus muscu	c 182	16.8	84.0	96916	10	AL670100	AL670100 Mouse DNA
110	16.8	84.0	22700	2	AC020792	AC020792 Mus muscu	c 183	16.8	84.0	96916	10	AL670100	AL670100 Mouse DNA
111	16.8	84.0	27688	2	AC146979	AC146979 Mus muscu	c 184	16.8	84.0	97325	10	AL928836	AL928836 Mouse DNA
112	16.8	84.0	27688	2	AC146979	AC146979 Mus muscu	c 185	16.8	84.0	97325	10	AL928836	AL928836 Mouse DNA
113	16.8	84.0	30210	2	AC100104	AC100104 Mus muscu	c 186	16.8	84.0	99268	9	AC091178	AC091178 Homo sapi
114	16.8	84.0	30210	2	AC100104	AC100104 Mus muscu	c 187	16.8	84.0	99268	9	AC091178	AC091178 Homo sapi
115	16.8	84.0	31638	1	RNEXRGN	RNEXRGN	c 188	16.8	84.0	99606	2	AP001871	AP001871 Homo sapi
116	16.8	84.0	31638	1	RNEXRGN	RNEXRGN	c 189	16.8	84.0	99606	2	AP001871	AP001871 Homo sapi
117	16.8	84.0	40693	10	BX784400	BX784400 Mouse DNA	c 190	16.8	84.0	100000	9	AP000508	AP000508 Homo sapi
118	16.8	84.0	40693	10	BX784400	BX784400 Mouse DNA	c 191	16.8	84.0	100803	10	AF367966	AF367966 Mus muscu
119	16.8	84.0	41506	9	AC004205	AC004205 Homo sapi	c 192	16.8	84.0	100803	10	AF367966	AF367966 Mus muscu
120	16.8	84.0	41506	9	AC004205	AC004205 Homo sapi	c 193	16.8	84.0	101563	10	AL807817	AL807817 Mouse DNA
121	16.8	84.0	42477	2	AC068805	AC068805 Mus muscu	c 194	16.8	84.0	101563	10	AL807817	AL807817 Mouse DNA
122	16.8	84.0	42477	2	AC068805	AC068805 Mus muscu	c 195	16.8	84.0	103842	10	AL840633	AL840633 Mouse DNA
123	16.8	84.0	42601	9	AC004210	AC004210 Homo sapi	c 196	16.8	84.0	103842	10	AL840633	AL840633 Mouse DNA
124	16.8	84.0	42601	9	AC004210	AC004210 Homo sapi	c 197	16.8	84.0	104509	10	AL928577	AL928577 Mouse DNA
125	16.8	84.0	43003	9	AL845556	AL845556 Human DNA	c 198	16.8	84.0	104509	10	AL928577	AL928577 Mouse DNA
126	16.8	84.0	43003	9	AL845556	AL845556 Human DNA	c 199	16.8	84.0	104746	2	AC022698	AC022698 Mus muscu
127	16.8	84.0	47355	10	AL672214	AL672214 Mouse DNA	c 200	16.8	84.0	104746	2	AC022698	AC022698 Mus muscu
128	16.8	84.0	47355	10	AL672214	AL672214 Mouse DNA	c 201	16.8	84.0	104892	10	AE008683_3	AE008683_3
129	16.8	84.0	49445	2	AC099873	AC099873 Mus muscu	c 202	16.8	84.0	104892	10	AE008683_3	AE008683_3
130	16.8	84.0	49445	2	AC099873	AC099873 Mus muscu	c 203	16.8	84.0	106635	2	AY555279	AY555279 Mus muscu
131	16.8	84.0	56934	2	AP000582	AP000582 Homo sapi	c 204	16.8	84.0	106635	2	AY555279	AY555279 Mus muscu
132	16.8	84.0	56934	2	AP000582	AP000582 Homo sapi	c 205	16.8	84.0	107898	10	AL627349	AL627349 Mouse DNA
133	16.8	84.0	59364	2	AC102347	AC102347 Mus muscu	c 206	16.8	84.0	107898	10	AL627349	AL627349 Mouse DNA
134	16.8	84.0	59364	2	AC102347	AC102347 Mus muscu	c 207	16.8	84.0	110000	2	RME591985_09	RME591985_09
135	16.8	84.0	60327	6	AX695572	AX695572 Sequence	c 208	16.8	84.0	110000	1	RME591985_09	RME591985_09
136	16.8	84.0	60327	6	AX695572	AX695572 Sequence	c 209	16.8	84.0	110000	2	AC101867_2	AC101867_2
137	16.8	84.0	61157	2	AC137956	AC137956 Mus muscu	c 210	16.8	84.0	110000	2	AC101867_2	AC101867_2
138	16.8	84.0	61157	2	AC137956	AC137956 Mus muscu	c 211	16.8	84.0	110000	2	AC101867_2	AC101867_2
139	16.8	84.0	62897	2	AC101230	AC101230 Mus muscu	c 212	16.8	84.0	110000	2	AL954691_1	AL954691_1
140	16.8	84.0	62897	2	AC101230	AC101230 Mus muscu	c 213	16.8	84.0	110000	2	AL954691_1	AL954691_1
141	16.8	84.0	62988	2	AC113449	AC113449 Mus muscu	c 214	16.8	84.0	110000	2	AX294366_2	AX294366_2
142	16.8	84.0	62988	2	AC113449	AC113449 Mus muscu	c 215	16.8	84.0	110000	10	AY36310252_1	AY36310252_1
143	16.8	84.0	63543	2	AC100544	AC100544 Mus muscu	c 216	16.8	84.0	110000	10	AY36310252_1	AY36310252_1
144	16.8	84.0	63543	2	AC100544	AC100544 Mus muscu	c 217	16.8	84.0	110000	10	AE008683_2	AE008683_2
145	16.8	84.0	64700	2	AC102043	AC102043 Mus muscu	c 218	16.8	84.0	110000	10	AE008683_2	AE008683_2
146	16.8	84.0	64700	2	AC102043	AC102043 Mus muscu	c 219	16.8	84.0	110892	9	AP000633	AP000633 Homo sapi
147	16.8	84.0	66116	2	AC116518	AC116518 Mus muscu	c 220	16.8	84.0	110892	9	AP000633	AP000633 Homo sapi
148	16.8	84.0	66116	2	AC116518	AC116518 Mus muscu	c 221	16.8	84.0	113581	2	AC109191	AC109191 Mus muscu
149	16.8	84.0	66929	2	AC121287	AC121287 Mus muscu	c 222	16.8	84.0	113581	2	AC109191	AC109191 Mus muscu
150	16.8	84.0	66929	2	AC121287	AC121287 Mus muscu	c 223	16.8	84.0	116320	9	AX248310	AX248310 Human DNA
151	16.8	84.0	68233	6	AX695956	AX695956 Sequence	c 224	16.8	84.0	116320	9	AX248310	AX248310 Human DNA
152	16.8	84.0	68233	6	AX695956	AX695956 Sequence	c 225	16.8	84.0	117000	10	AC087871	AC087871 Genomic s
153	16.8	84.0	69440	2	AC118034	AC118034 Mus muscu	c 226	16.8	84.0	117000	10	AC087871	AC087871 Genomic s
154	16.8	84.0	69440	2	AC118034	AC118034 Mus muscu	c 227	16.8	84.0	120197	9	AL671883	AL671883 Human DNA
155	16.8	84.0	70137	10	AL713988	AL713988 Mouse DNA	c 228	16.8	84.0	120197	9	AL671883	AL671883 Human DNA
156	16.8	84.0	70137	10	AL713988	AL713988 Mouse DNA	c 229	16.8	84.0	124933	10	AF545858	AF545858 Mus muscu
157	16.8	84.0	70194	2	AC133888	AC133888 Mus muscu	c 230	16.8	84.0	124933	10	AF545858	AF545858 Mus muscu
158	16.8	84.0	70194	2	AC133888	AC133888 Mus muscu	c 231	16.8	84.0	125641	10	AL589742	AL589742 Mouse DNA
159	16.8	84.0	70991	10	AY225417	AY225417 Mus muscu	c 232	16.8	84.0	125641	10	AL589742	AL589742 Mouse DNA
160	16.8	84.0	70991	10	AY225417	AY225417 Mus muscu	c 233	16.8	84.0	127025	2	AC125522	AC125522 Mus muscu
161	16.8	84.0	75039	2	AC102185	AC102185 Mus muscu	c 234	16.8	84.0	127025	2	AC125522	AC125522 Mus muscu
162	16.8	84.0	75039	2	AC102185	AC102185 Mus muscu	c 235	16.8	84.0	127380	10	AC125113	AC125113 Mus muscu
163	16.8	84.0	78135	2	AC139059	AC139059 Mus muscu	c 236	16.8	84.0	127380	10	AC125113	AC125113 Mus muscu
164	16.8	84.0	78135	2	AC139059	AC139059 Mus muscu	c 237	16.8	84.0	129451	2	AC073783	AC073783 Mus muscu
165	16.8	84.0	79925	9	AC093719	AC093719 Homo sapi	c 238	16.8	84.0	129451	2	AC073783	AC073783 Mus muscu

C 239	16.8	84.0	130301	10	EX813328	Mouse DNA	C 312	16.8	84.0	150643	2	AP001844	Homo sapi
C 240	16.8	84.0	130301	10	EX813328	Mouse DNA	C 313	16.8	84.0	151639	2	AP001193	Homo sapi
C 241	16.8	84.0	130474	2	AC079039	Mus muscu	C 314	16.8	84.0	151639	2	AP001193	Homo sapi
C 242	16.8	84.0	130474	2	AC079039	Mus muscu	C 315	16.8	84.0	152275	2	AC147129	Mus muscu
C 243	16.8	84.0	131732	10	AF259071	Mus muscu	C 316	16.8	84.0	152275	2	AC147129	Mus muscu
C 244	16.8	84.0	131732	10	AF259071	Mus muscu	C 317	16.8	84.0	153847	10	AL591763	Mouse DNA
C 245	16.8	84.0	132342	10	AC132566	Mus muscu	C 318	16.8	84.0	153847	10	AL591763	Mouse DNA
C 246	16.8	84.0	132342	10	AC132566	Mus muscu	C 319	16.8	84.0	154373	10	AC121988	Mus muscu
C 247	16.8	84.0	132668	2	AC099829	Homo sapi	C 320	16.8	84.0	154373	10	AC121988	Mus muscu
C 248	16.8	84.0	132668	2	AC099829	Homo sapi	C 321	16.8	84.0	154754	2	AC105982	Mus muscu
C 249	16.8	84.0	132688	2	AC119878	Mus muscu	C 322	16.8	84.0	154754	2	AC105982	Mus muscu
C 250	16.8	84.0	132688	2	AC119878	Mus muscu	C 323	16.8	84.0	155954	10	AL928860	Mouse DNA
C 251	16.8	84.0	135432	2	AL645589	Mus muscu	C 324	16.8	84.0	155954	10	AL928860	Mouse DNA
C 252	16.8	84.0	135432	2	AL645589	Mus muscu	C 325	16.8	84.0	156468	10	AL954817	Mouse DNA
C 253	16.8	84.0	135453	2	AC118736	Mus muscu	C 326	16.8	84.0	156468	10	AL954817	Mouse DNA
C 254	16.8	84.0	135453	2	AC118736	Mus muscu	C 327	16.8	84.0	156603	10	AC120787	Mus muscu
C 255	16.8	84.0	136155	2	AC132586	Mus muscu	C 328	16.8	84.0	156603	10	AC120787	Mus muscu
C 256	16.8	84.0	136155	10	AC132586	Mus muscu	C 329	16.8	84.0	157019	10	AL772156	Mouse DNA
C 257	16.8	84.0	137909	2	AC117640	Mus muscu	C 330	16.8	84.0	157019	10	AL772156	Mouse DNA
C 258	16.8	84.0	137909	2	AC117640	Mus muscu	C 331	16.8	84.0	157055	10	AC134524	Mus muscu
C 259	16.8	84.0	139160	2	AC131230	Mus muscu	C 332	16.8	84.0	157055	10	AC134524	Mus muscu
C 260	16.8	84.0	139160	2	AC131230	Mus muscu	C 333	16.8	84.0	157958	10	AC147983	Mus muscu
C 261	16.8	84.0	140494	10	BX276179	Mouse DNA	C 334	16.8	84.0	157958	10	AC147983	Mus muscu
C 262	16.8	84.0	140494	10	BX276179	Mouse DNA	C 335	16.8	84.0	158006	10	AC127687	Mus muscu
C 263	16.8	84.0	141036	2	AC023807	Mus muscu	C 336	16.8	84.0	158006	10	AC127687	Mus muscu
C 264	16.8	84.0	141036	2	AC023807	Mus muscu	C 337	16.8	84.0	158009	10	AL837508	Mouse DNA
C 265	16.8	84.0	141324	2	AL645466	Mouse DNA	C 338	16.8	84.0	158009	10	AL837508	Mouse DNA
C 266	16.8	84.0	141324	10	AL645466	Mouse DNA	C 339	16.8	84.0	158764	10	AC132083	Mus muscu
C 267	16.8	84.0	141678	10	AC102722	Mus muscu	C 340	16.8	84.0	158764	10	AC132083	Mus muscu
C 268	16.8	84.0	141678	10	AC102722	Mus muscu	C 341	16.8	84.0	159669	2	AC102017	Mus muscu
C 269	16.8	84.0	142565	10	AL772205	Mouse DNA	C 342	16.8	84.0	159669	2	AC102017	Mus muscu
C 270	16.8	84.0	142565	10	AL772205	Mouse DNA	C 343	16.8	84.0	159731	10	AL773512	Mouse DNA
C 271	16.8	84.0	143342	10	AL591864	Mouse DNA	C 344	16.8	84.0	159731	10	AL773512	Mouse DNA
C 272	16.8	84.0	143342	10	AL591864	Mouse DNA	C 345	16.8	84.0	160015	10	AC141873	Mus muscu
C 273	16.8	84.0	143875	10	AC101709	Mus muscu	C 346	16.8	84.0	160015	10	AC141873	Mus muscu
C 274	16.8	84.0	143875	10	AC101709	Mus muscu	C 347	16.8	84.0	160180	10	AC129016	Mus muscu
C 275	16.8	84.0	143963	10	AL844176	Mouse DNA	C 348	16.8	84.0	160180	10	AC129016	Mus muscu
C 276	16.8	84.0	143963	10	AL844176	Mouse DNA	C 349	16.8	84.0	160373	10	AC129291	Mus muscu
C 277	16.8	84.0	144116	2	AC105959	Mus muscu	C 350	16.8	84.0	160373	10	AC129291	Mus muscu
C 278	16.8	84.0	144116	2	AC105959	Mus muscu	C 351	16.8	84.0	160449	10	AC130549	Mus muscu
C 279	16.8	84.0	144339	10	AL627347	Mouse DNA	C 352	16.8	84.0	160449	10	AC130549	Mus muscu
C 280	16.8	84.0	144339	10	AL627347	Mouse DNA	C 353	16.8	84.0	160728	10	AC124432	Mouse DNA
C 281	16.8	84.0	145273	2	AC133949	Mus muscu	C 354	16.8	84.0	160728	10	AC124432	Mouse DNA
C 282	16.8	84.0	145273	2	AC133949	Mus muscu	C 355	16.8	84.0	160869	2	AC026378	Mus muscu
C 283	16.8	84.0	146015	10	AC127545	Mus muscu	C 356	16.8	84.0	160869	2	AC026378	Mus muscu
C 284	16.8	84.0	146015	10	AC127545	Mus muscu	C 357	16.8	84.0	161335	2	AY510701	Mus muscu
C 285	16.8	84.0	146175	10	AL663084	Mouse DNA	C 358	16.8	84.0	161335	2	AY510701	Mus muscu
C 286	16.8	84.0	146175	10	AL663084	Mouse DNA	C 359	16.8	84.0	161492	2	AC131772	Mus muscu
C 287	16.8	84.0	146301	9	AC114494	Mouse DNA	C 360	16.8	84.0	161492	2	AC131772	Mus muscu
C 288	16.8	84.0	146301	9	AC114494	Mouse DNA	C 361	16.8	84.0	161940	2	AC079289	Homo sapi
C 289	16.8	84.0	146420	2	AP001140	Homo sapi	C 362	16.8	84.0	161940	2	AC079289	Homo sapi
C 290	16.8	84.0	146420	2	AP001140	Homo sapi	C 363	16.8	84.0	162002	2	AC138741	Mus muscu
C 291	16.8	84.0	146527	10	EX088584	Mouse DNA	C 364	16.8	84.0	162002	2	AC138741	Mus muscu
C 292	16.8	84.0	146527	10	EX088584	Mouse DNA	C 365	16.8	84.0	162919	2	AC102687	Mus muscu
C 293	16.8	84.0	146547	10	AL929471	Mouse DNA	C 366	16.8	84.0	162919	2	AC102687	Mus muscu
C 294	16.8	84.0	146547	10	AL929471	Mouse DNA	C 367	16.8	84.0	162996	2	AC148174	Mus muscu
C 295	16.8	84.0	147984	10	AF36796753	Mus muscu	C 368	16.8	84.0	162996	2	AC148174	Mus muscu
C 296	16.8	84.0	147984	10	AF36796753	Mus muscu	C 369	16.8	84.0	163240	10	AL935150	Mouse DNA
C 297	16.8	84.0	148076	9	AL845443	Human DNA	C 370	16.8	84.0	163240	10	AL935150	Mouse DNA
C 298	16.8	84.0	148076	9	AL845443	Human DNA	C 371	16.8	84.0	163287	10	AC122290	Mus muscu
C 299	16.8	84.0	148263	10	AC127573	Mus muscu	C 372	16.8	84.0	163287	10	AC122290	Mus muscu
C 300	16.8	84.0	148263	10	AC127573	Mus muscu	C 373	16.8	84.0	163306	2	AC119937	Mus muscu
C 301	16.8	84.0	148334	10	AY555278	Mus muscu	C 374	16.8	84.0	163306	2	AC119937	Mus muscu
C 302	16.8	84.0	148334	10	AY555278	Mus muscu	C 375	16.8	84.0	163764	2	AC121099	Mus muscu
C 303	16.8	84.0	148447	10	AC125141	Mus muscu	C 376	16.8	84.0	163764	2	AC121099	Mus muscu
C 304	16.8	84.0	148447	10	AC125141	Mus muscu	C 377	16.8	84.0	163969	10	AC114605	Mus muscu
C 305	16.8	84.0	149368	2	AC124395	Mus muscu	C 378	16.8	84.0	163969	10	AC114605	Mus muscu
C 306	16.8	84.0	149368	2	AC124395	Mus muscu	C 379	16.8	84.0	164223	10	AC127228	Mus muscu
C 307	16.8	84.0	149451	10	AC115294	Mus muscu	C 380	16.8	84.0	164223	10	AC127228	Mus muscu
C 308	16.8	84.0	149451	10	AC115294	Mus muscu	C 381	16.8	84.0	164311	10	AC115881	Mus muscu
C 309	16.8	84.0	150235	5	AL590152	zebrafish	C 382	16.8	84.0	164311	10	AC115881	Mus muscu
C 310	16.8	84.0	150235	5	AL590152	zebrafish	C 383	16.8	84.0	164443	10	AL672285	Mouse DNA
C 311	16.8	84.0	150643	2	AP001844	Homo sapi	C 384	16.8	84.0	164443	10	AL672285	Mouse DNA

385	16.8	84.0	164866	2	AC091102	AC091102 Homo sapi
386	16.8	84.0	164866	2	AC091102	AC091102 Homo sapi
387	16.8	84.0	165703	10	AC139296	AC139296 Mus muscu
388	16.8	84.0	165703	10	AC139296	AC139296 Mus muscu
389	16.8	84.0	165733	2	AC107706	AC107706 Mus muscu
390	16.8	84.0	165733	2	AC107706	AC107706 Mus muscu
391	16.8	84.0	166042	9	AP006289	AP006289 Homo sapi
392	16.8	84.0	166042	9	AP006289	AP006289 Homo sapi
393	16.8	84.0	166231	2	AC087037	AC087037 Mus muscu
394	16.8	84.0	166231	2	AC087037	AC087037 Mus muscu
395	16.8	84.0	166332	2	AC135667	AC135667 Mus muscu
396	16.8	84.0	166332	2	AC135667	AC135667 Mus muscu
397	16.8	84.0	166335	2	AC073799	AC073799 Mus muscu
398	16.8	84.0	166335	2	AC073799	AC073799 Mus muscu
399	16.8	84.0	167166	10	AL733520	AL733520 Mouse DNA
400	16.8	84.0	167166	10	AL733520	AL733520 Mouse DNA
401	16.8	84.0	167418	10	AL845294	AL845294 Mouse DNA
402	16.8	84.0	167418	10	AL845294	AL845294 Mouse DNA
403	16.8	84.0	168105	2	AC115069	AC115069 Mus muscu
404	16.8	84.0	168105	2	AC115069	AC115069 Mus muscu
405	16.8	84.0	168117	10	AC133352	AC133352 Mus muscu
406	16.8	84.0	168117	10	AC133352	AC133352 Mus muscu
407	16.8	84.0	168251	2	AC112940	AC112940 Mus muscu
408	16.8	84.0	168251	2	AC112940	AC112940 Mus muscu
409	16.8	84.0	168515	10	AC130215	AC130215 Mus muscu
410	16.8	84.0	168515	10	AC130215	AC130215 Mus muscu
411	16.8	84.0	168698	2	AC127583	AC127583 Mus muscu
412	16.8	84.0	168698	2	AC127583	AC127583 Mus muscu
413	16.8	84.0	168916	2	AC124822	AC124822 Mus muscu
414	16.8	84.0	168916	2	AC124822	AC124822 Mus muscu
415	16.8	84.0	169247	10	AC103621	AC103621 Mus muscu
416	16.8	84.0	169247	10	AC103621	AC103621 Mus muscu
417	16.8	84.0	169481	10	AC122403	AC122403 Mus muscu
418	16.8	84.0	169481	10	AC122403	AC122403 Mus muscu
419	16.8	84.0	169815	10	AC120057	AC120057 Mus muscu
420	16.8	84.0	169815	10	AC120057	AC120057 Mus muscu
421	16.8	84.0	170020	9	EX072579	EX072579 Human DNA
422	16.8	84.0	170020	9	EX072579	EX072579 Human DNA
423	16.8	84.0	170119	2	AC143515	AC143515 Macaca mu
424	16.8	84.0	170119	2	AC143515	AC143515 Macaca mu
425	16.8	84.0	170369	2	AC102188	AC102188 Mus muscu
426	16.8	84.0	170369	2	AC102188	AC102188 Mus muscu
427	16.8	84.0	170548	10	AC074041	AC074041 Mus muscu
428	16.8	84.0	170548	10	AC074041	AC074041 Mus muscu
429	16.8	84.0	171122	10	AC124394	AC124394 Mus muscu
430	16.8	84.0	171122	10	AC124394	AC124394 Mus muscu
431	16.8	84.0	171209	10	AC132622	AC132622 Mus muscu
432	16.8	84.0	171209	10	AC132622	AC132622 Mus muscu
433	16.8	84.0	171261	10	AC110514	AC110514 Mus muscu
434	16.8	84.0	171261	10	AC110514	AC110514 Mus muscu
435	16.8	84.0	171527	2	AC129942	AC129942 Mus muscu
436	16.8	84.0	171527	2	AC129942	AC129942 Mus muscu
437	16.8	84.0	172237	10	AC134475	AC134475 Mus muscu
438	16.8	84.0	172237	10	AC134475	AC134475 Mus muscu
439	16.8	84.0	172263	2	AC119280	AC119280 Mus muscu
440	16.8	84.0	172263	2	AC119280	AC119280 Mus muscu
441	16.8	84.0	172692	5	EX005380	EX005380 Zebrafish
442	16.8	84.0	172692	5	EX005380	EX005380 Zebrafish
443	16.8	84.0	172694	2	AC124116	AC124116 Mus muscu
444	16.8	84.0	172694	2	AC124116	AC124116 Mus muscu
445	16.8	84.0	173181	10	AC145525	AC145525 Mus muscu
446	16.8	84.0	173181	10	AC145525	AC145525 Mus muscu
447	16.8	84.0	173351	10	AL732502	AL732502 Mouse DNA
448	16.8	84.0	173351	10	AL732502	AL732502 Mouse DNA
449	16.8	84.0	173769	10	AL669819	AL669819 Mouse DNA
450	16.8	84.0	173769	10	AL669819	AL669819 Mouse DNA
451	16.8	84.0	173926	2	MM454K24	MM454K24 Mus muscu
452	16.8	84.0	173926	2	MM454K24	MM454K24 Mus muscu
453	16.8	84.0	174670	2	AC112933	AC112933 Mus muscu
454	16.8	84.0	174670	2	AC112933	AC112933 Mus muscu
455	16.8	84.0	174750	10	AC124359	AC124359 Mus muscu
456	16.8	84.0	174750	10	AC124359	AC124359 Mus muscu
457	16.8	84.0	174928	10	AC121975	AC121975 Mus muscu

458	16.8	84.0	174928	10	AC121975	AC121975 Mus muscu
459	16.8	84.0	175092	10	AC122491	AC122491 Mus muscu
460	16.8	84.0	175092	10	AC122491	AC122491 Mus muscu
461	16.8	84.0	175131	10	AC121821	AC121821 Mus muscu
462	16.8	84.0	175131	10	AC121821	AC121821 Mus muscu
463	16.8	84.0	175316	2	AC136516	AC136516 Mus muscu
464	16.8	84.0	175316	2	AC136516	AC136516 Mus muscu
465	16.8	84.0	175340	2	AC110159	AC110159 Mus muscu
466	16.8	84.0	175340	2	AC110159	AC110159 Mus muscu
467	16.8	84.0	175809	10	AL714027	AL714027 Mouse DNA
468	16.8	84.0	175809	10	AL714027	AL714027 Mouse DNA
469	16.8	84.0	175878	2	AC122554	AC122554 Mus muscu
470	16.8	84.0	175878	2	AC122554	AC122554 Mus muscu
471	16.8	84.0	175987	10	AC079273	AC079273 Mus muscu
472	16.8	84.0	175987	10	AC079273	AC079273 Mus muscu
473	16.8	84.0	176056	10	AC073437	AC073437 Mus muscu
474	16.8	84.0	176056	10	AC073437	AC073437 Mus muscu
475	16.8	84.0	176278	10	AC126557	AC126557 Mus muscu
476	16.8	84.0	176278	10	AC126557	AC126557 Mus muscu
477	16.8	84.0	176284	9	AB054536	AB054536 Pan trogl
478	16.8	84.0	176284	9	AB054536	AB054536 Pan trogl
479	16.8	84.0	176304	2	AC102182	AC102182 Mus muscu
480	16.8	84.0	176304	2	AC102182	AC102182 Mus muscu
481	16.8	84.0	176436	10	AC129937	AC129937 Mus muscu
482	16.8	84.0	176436	10	AC129937	AC129937 Mus muscu
483	16.8	84.0	176541	10	AC127367	AC127367 Mus muscu
484	16.8	84.0	176541	10	AC127367	AC127367 Mus muscu
485	16.8	84.0	176711	10	AC137969	AC137969 Mus muscu
486	16.8	84.0	176711	10	AC137969	AC137969 Mus muscu
487	16.8	84.0	176871	10	AC124687	AC124687 Mus muscu
488	16.8	84.0	176871	10	AC124687	AC124687 Mus muscu
489	16.8	84.0	177183	2	AC112973	AC112973 Mus muscu
490	16.8	84.0	177183	2	AC112973	AC112973 Mus muscu
491	16.8	84.0	177256	10	AL691476	AL691476 Mouse DNA
492	16.8	84.0	177256	10	AL691476	AL691476 Mouse DNA
493	16.8	84.0	177302	10	AC127415	AC127415 Mus muscu
494	16.8	84.0	177302	10	AC127415	AC127415 Mus muscu
495	16.8	84.0	177450	2	AC103674	AC103674 Mus muscu
496	16.8	84.0	177450	2	AC103674	AC103674 Mus muscu
497	16.8	84.0	177788	2	AC069310	AC069310 Mus muscu
498	16.8	84.0	177788	2	AC069310	AC069310 Mus muscu
499	16.8	84.0	177955	2	AC131034	AC131034 Mus muscu
500	16.8	84.0	177955	2	AC131034	AC131034 Mus muscu

## ALIGNMENTS

RESULT 1  
LOCUS AR093067 20 bp DNA  
DEFINITION Sequence 162 from patent US 5998383.  
ACCESSION AR093067  
VERSION AR093067.1 GI:10019819  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 20)  
AUTHORS Wright,J.A. and Young,A.H.  
TITL Antitumor antisense sequences directed against ribonucleotide  
reductase  
JOURNAL Patent: US 5998383-A 162 07-DEC-1999;  
FEATURES Location/Qualifiers  
source 1..20  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy 1 GGACATGCCCGGCATGTCC 20  
Db 1 GGACATGCCCGGCATGTCC 20

RESULT 2  
AR093067/c 20 bp DNA linear PAT 08-SEP-2000  
LOCUS Sequence 162 from patent US 5998383.  
DEFINITION AR093067  
ACCESSION AR093067.1 GI:10019819  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .20  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20  
Db 20 GGACATGCCCGGCATGTCC 1

RESULT 3  
E32518 20 bp DNA linear PAT 18-JUN-2001  
LOCUS Scavenger receptor-like protein.  
DEFINITION E32518  
ACCESSION E32518.1 GI:13026765  
VERSION JP 1999123094-A/18.  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20  
Db 1 GGACATGCCCGGCATGTCC 20

RESULT 4  
E32518/c 20 bp DNA linear PAT 18-JUN-2001  
LOCUS Scavenger receptor-like protein.  
DEFINITION E32518  
ACCESSION E32518.1 GI:13026765  
VERSION JP 1999123094-A/18.  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20  
Db 20 GGACATGCCCGGCATGTCC 1

RESULT 5  
AR359569 20 bp DNA linear PAT 17-AUG-2003  
LOCUS Sequence 162 from patent US 6593305.  
DEFINITION AR359569  
ACCESSION AR359569.1 GI:33766292  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .20  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20  
 DB 1 GGACATGCCCCGGCATGTCC 20

RESULT 6  
 AR359569/c  
 LOCUS AR359569 20 bp DNA linear PAT 17-AUG-2003  
 DEFINITION Sequence 162 from patent US 6593305.  
 ACCESSION AR359569  
 VERSION AR359569.1 GI:33766292  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 20)  
 AUTHORS Wright,J.A.  
 TITLE Antitumor antisense sequences directed against R1 and R2 components of ribonucleotide reductase  
 JOURNAL Patent: US 6593305-A 162 15-UTL-2003;  
 FEATURES Location/Qualifiers  
 source 1..20  
 /organism="Unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20  
 DB 20 GGACATGCCCCGGCATGTCC 1

RESULT 7  
 AX339212  
 LOCUS AX339212 20 bp DNA linear PAT 10-JAN-2002  
 DEFINITION Sequence 6 from Patent WO0196602.  
 ACCESSION AX339212  
 VERSION AX339212.1 GI:18135473  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1  
 AUTHORS Yang,A.L. and Festing,M.  
 TITLE Methods and materials to determine the p53 status of a sample by determining the binding of p53 to a vector  
 JOURNAL Patent: WO 0196602-A 6 20-DEC-2001;  
 FEATURES MEDICAL RESEARCH COUNCIL (GB)  
 source Location/Qualifiers  
 1..20  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Consensus sequence"

ORIGIN  
 Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20  
 DB 1 GGACATGCCCCGGCATGTCC 20

RESULT 8

AX339212/c  
 LOCUS AX339212 20 bp DNA linear PAT 10-JAN-2002  
 DEFINITION Sequence 6 from Patent WO0196602.  
 ACCESSION AX339212  
 VERSION AX339212.1 GI:18135473  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1  
 AUTHORS Yang,A.L. and Festing,M.  
 TITLE Methods and materials to determine the p53 status of a sample by determining the binding of p53 to a vector  
 JOURNAL Patent: WO 0196602-A 6 20-DEC-2001;  
 FEATURES MEDICAL RESEARCH COUNCIL (GB)  
 source Location/Qualifiers  
 1..20  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Consensus sequence"

ORIGIN  
 Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20  
 DB 20 GGACATGCCCCGGCATGTCC 1

RESULT 9  
 HSDMP53  
 LOCUS HSDMP53 20 bp DNA linear PRI 05-JUN-1992  
 DEFINITION H.sapiens DNA binding motif recognized by p53 protein containing complexes.  
 ACCESSION X63571  
 VERSION X63571.1 GI:30483  
 KEYWORDS DNA motif; high affinity DNA binding motif.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 20)  
 AUTHORS Funk,W.D., Pak,D.T., Karas,R.H., Wright,W.E. and Shay,J.W.  
 TITLE A transcriptionally active DNA-binding site for human p53 protein complexes  
 JOURNAL Mol. Cell. Biol. 12 (6), 2866-2871 (1992)  
 MEDLINE 92269860  
 PUBMED 1588974  
 REFERENCE 2 (bases 1 to 20)  
 AUTHORS Shay,J.W.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-JAN-1992) J.W. Shay, The University of Texas, Southwestern Medical Center at Dallas, Dept of Cell Biology & Neuroscience, 5323 Harry Hines Boulevard, Dallas TX 75235-9039, USA

COMMENT  
 See also X63905-7 & S.E. Kern et al, Science 252:1708-1711(1991).  
 This sequence is a high-affinity DNA binding motif recognized by p53 protein containing complexes and is likely to be a regulatory signal. Variations of this sequence is found in the 5' flanking region, 3' flanking or untranslated regions of many genes associated with cellular proliferation including a perfect match in the 5' flanking regions of the c-Ha-ras-1 oncogene, the gene for DNA ligase 1 & in the 3' untranslated region of the tyrosine kinase receptor gene. When placed upstream of a reporter construct, this consensus sequence promotes p53 dependent transcription.

FEATURES  
 source Location/Qualifiers  
 1..20  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

ORIGIN /cell\_type="diploid fibroblasts"

Query Match 100.0%; Score 20; DB 9; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 1 GGACATGCCCGGCATGTCC 20

Db 1 GGACATGCCCGGCATGTCC 20

RESULT 10  
 HSDMP53/c  
 LOCUS  
 DEFINITION 20 bp DNA linear PRI 05-JUN-1992  
 H. sapiens DNA binding motif recognized by p53 protein containing  
 complexes.

ACCESSION X63571.1 GI:30483  
 VERSION X63571.1  
 KEYWORDS DNA motif; high affinity DNA binding motif.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 20)  
 Funk, W.D., Pak, D.T., Karas, R.H., Wright, W.E. and Shay, J.W.  
 A transcriptionally active DNA-binding site for human p53 protein  
 complexes  
 Mol. Cell. Biol. 12 (6), 2866-2871 (1992)

JOURNAL MEDLINE  
 PUBMED 92269860  
 1588974  
 2 (bases 1 to 20)  
 Shay, J.W.  
 Direct Submission  
 Submitted (09-JUN-1992) J.W. Shay, The University of Texas,  
 Southwestern Medical Center at Dallas, Dept of Cell Biology &  
 Neuroscience, 5323 Harry Hines Boulevard, Dallas TX 75235-9039,  
 USA

COMMENT See also X63905-7 & S.R. Kern et al. Science 252:1708-1711 (1991).  
 This sequence is a high-affinity DNA binding motif recognized by  
 p53 protein containing complexes and is likely to be a regulatory  
 signal. Variations of this sequence is found in the 5' flanking,  
 15', 3' flanking or untranslated regions of many genes associated  
 with cellular proliferation including a perfect match in the 5'  
 flanking regions of the c-Ha-ras-1 oncogene, the gene for DNA  
 ligase I & in the 3' untranslated region of the tyrosine kinase  
 receptor gene. When placed upstream of a reporter construct, this  
 consensus sequence promotes p53 dependent transcription.

FEATURES  
 source  
 1..20  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /cell\_type="diploid fibroblasts"

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 1 GGACATGCCCGGCATGTCC 20

Db 20 GGACATGCCCGGCATGTCC 1

RESULT 11  
 AX472546 24 bp DNA linear PAT 09-AUG-2002  
 LOCUS  
 DEFINITION Sequence 41 from Patent WO02052039.  
 ACCESSION AX472546  
 VERSION AX472546.1 GI:22207450  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 synthetic construct

ORGANISM synthetic construct  
 artificial sequences.

REFERENCE 1  
 AUTHORS Blais, Y., Rousseau, P., Leblanc, B. and Camero, R.N.  
 TITLE Methods for selecting and producing selective pharmaceutical  
 compounds and compositions using an established genetically altered  
 cell-based library responsive to transcription factors; genetic  
 constructs and library therefor  
 JOURNAL Patent: WO 02052039-A 41 04-JUL-2002;  
 Geneka Biotechnology Inc. (CA)  
 FEATURES  
 source  
 1..24  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Oligonucleotide"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 5 GGACATGCCCGGCATGTCC 24

Db 5 GGACATGCCCGGCATGTCC 24

RESULT 12  
 AX472546 24 bp DNA linear PAT 09-AUG-2002  
 LOCUS  
 DEFINITION Sequence 41 from Patent WO02052039.  
 ACCESSION AX472546  
 VERSION AX472546.1 GI:22207450  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 synthetic construct  
 synthetic construct  
 artificial sequences.

AUTHORS Blais, Y., Rousseau, P., Leblanc, B. and Camero, R.N.  
 TITLE Methods for selecting and producing selective pharmaceutical  
 compounds and compositions using an established genetically altered  
 cell-based library responsive to transcription factors; genetic  
 constructs and library therefor  
 JOURNAL Patent: WO 02052039-A 41 04-JUL-2002;  
 Geneka Biotechnology Inc. (CA)  
 FEATURES  
 source  
 1..24  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Oligonucleotide"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 24 GGACATGCCCGGCATGTCC 5

Db 24 GGACATGCCCGGCATGTCC 5

RESULT 13  
 AX476842 24 bp DNA linear PAT 12-AUG-2002  
 LOCUS  
 DEFINITION Sequence 19 from Patent WO02052037.  
 ACCESSION AX476842  
 VERSION AX476842.1 GI:22216118  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 synthetic construct  
 synthetic construct  
 artificial sequences.

AUTHORS Larose,A.M., Rousseau,P., Lebanc,B. and Camato,R.  
 TITLE Method for screening and/or identifying factors that bind to  
 nucleic acids  
 JOURNAL Patent: WO 02052037-A 19 04-JUL-2002;  
 Geneka Biotechnology Inc. (CA)

## FEATURES

source  
 1. .24  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="NABE-probes"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 5 GGACATGCCCGGCATGTCC 24

## RESULT 14

LOCUS AX476842/c 24 bp DNA linear PAT 12-AUG-2002  
 DEFINITION Sequence 19 from Patent WO02052037.  
 ACCESSION AX476842  
 VERSION AX476842.1 GI:22216118  
 KEYWORDS

## SOURCE

synthetic construct  
 synthetic construct  
 artificial sequences.

## REFERENCE

1  
 AUTHORS Larose,A.M., Rousseau,P., Lebanc,B. and Camato,R.  
 TITLE Method for screening and/or identifying factors that bind to  
 nucleic acids  
 JOURNAL Patent: WO 02052037-A 19 04-JUL-2002;  
 Geneka Biotechnology Inc. (CA)

## FEATURES

source  
 1. .24  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="NABE-probes"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 24 GGACATGCCCGGCATGTCC 5

## RESULT 15

LOCUS AX476861 24 bp DNA linear PAT 12-AUG-2002  
 DEFINITION Sequence 38 from Patent WO02052037.  
 ACCESSION AX476861  
 VERSION AX476861.1 GI:22216137  
 KEYWORDS

## SOURCE

synthetic construct  
 synthetic construct  
 artificial sequences.

## REFERENCE

1  
 AUTHORS Larose,A.M., Rousseau,P., Lebanc,B. and Camato,R.  
 TITLE Method for screening and/or identifying factors that bind to  
 nucleic acids  
 JOURNAL Patent: WO 02052037-A 38 04-JUL-2002;  
 Geneka Biotechnology Inc. (CA)

## FEATURES

source  
 1. .24

/organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Double stranded NABE"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 5 GGACATGCCCGGCATGTCC 24

RESULT 16  
 LOCUS AX476861/c 24 bp DNA linear PAT 12-AUG-2002  
 DEFINITION Sequence 38 from Patent WO02052037.  
 ACCESSION AX476861  
 VERSION AX476861.1 GI:22216137  
 KEYWORDS

## SOURCE

synthetic construct  
 synthetic construct  
 artificial sequences.

## REFERENCE

1  
 AUTHORS Larose,A.M., Rousseau,P., Lebanc,B. and Camato,R.  
 TITLE Method for screening and/or identifying factors that bind to  
 nucleic acids  
 JOURNAL Patent: WO 02052037-A 38 04-JUL-2002;  
 Geneka Biotechnology Inc. (CA)

## FEATURES

source  
 1. .24  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Double stranded NABE"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 24 GGACATGCCCGGCATGTCC 5

## RESULT 17

LOCUS BD272327 56 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Selectively replicating viral vector.  
 ACCESSION BD272327  
 VERSION BD272327.1 GI:33082095  
 KEYWORDS JP 2002541761-A/9.

## SOURCE

synthetic construct  
 synthetic construct  
 artificial sequences.

## REFERENCE

1 (bases 1 to 56)  
 AUTHORS Ramachandra,M. and Shabram,P.W.  
 TITLE Selectively replicating viral vector  
 JOURNAL Patent: JP 2002541761-A 9 10-DEC-2002;  
 CANUI INC

## COMMENT

CS Artificial Sequence  
 PN JP 2002541761-A/9  
 PD 10-DEC-2002 JP 2000576027  
 PR 14-OCT-1999 JP 09/172686  
 PR 15-OCT-1998 US 09/172686  
 PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM  
 PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,C12N5/10,C12N7/00, PC  
 C12N15/00  
 PC C12N5/00  
 CC complementary 5'-phosphorylated oligonucleotide containing CC

FEATURES  
Source  
CC binding sites(p53CON)  
FH Key Location/Qualifiers  
FT source 1..56  
Location/Qualifiers  
1..56  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 56;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
|||||  
7 GGACATGCCCGGCGCATGTCC 26

RESULT 18  
BD272327/c 56 bp DNA linear PAT 17-JUL-2003  
DEFINITION Selectively replicating viral vector.  
ACCESSION BD272327.1 GI:33082095  
KEYWORDS UP 2002541761-A/9.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
1 (bases 1 to 56)  
REFERENCE Ramachandra,M. and Shabram,P.W.  
AUTHORS Selectively replicating viral vector  
TITLE Patent: JP 2002541761-A 9 10-DEC-2002;  
JOURNAL CANJI INC

COMMENT  
OS Artificial Sequence  
PN JP 2002541761-A/9  
PD 10-DEC-2002  
PF 14-OCT-1999 JP 2000576027  
PR 15-OCT-1998 US 09/172686  
PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM  
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,C12N5/10,C12N7/00, PC  
C12N15/00,  
PC C12N5/00  
CC complementary 5'-phosphorylated oligonucleotide containing CC  
consensus p53  
CC binding sites(p53CON)  
FH Key Location/Qualifiers  
FT source 1..56  
Location/Qualifiers  
1..56  
/organism="Artificial Sequence".

FEATURES  
Source  
1..56  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 56;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
|||||  
51 GGACATGCCCGGCGCATGTCC 32

RESULT 19  
BD272328 56 bp DNA linear PAT 17-JUL-2003  
DEFINITION Selectively replicating viral vector.  
ACCESSION BD272328.1 GI:33082096  
VERSION

KEYWORDS JP 2002541761-A/10.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 56)  
AUTHORS Ramachandra,M. and Shabram,P.W.  
TITLE Selectively replicating viral vector  
JOURNAL Patent: JP 2002541761-A 10 10-DEC-2002;  
CANJI INC

COMMENT  
OS Artificial Sequence  
PN JP 2002541761-A/10  
PD 10-DEC-2002  
PF 14-OCT-1999 JP 2000576027  
PR 15-OCT-1998 US 09/172686  
PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM  
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,C12N5/10,C12N7/00, PC  
C12N15/00,  
PC C12N5/00  
CC complementary 5'-phosphorylated oligonucleotide containing CC  
consensus p53  
CC binding sites(p53CON)  
FH Key Location/Qualifiers  
FT source 1..56  
Location/Qualifiers  
1..56  
/organism="Artificial Sequence".

FEATURES  
Source  
1..56  
/organism="synthetic construct"

QY 1 GGACATGCCCGGCGCATGTCC 20  
|||||  
2 GGACATGCCCGGCGCATGTCC 21

RESULT 20  
BD272328 56 bp DNA linear PAT 17-JUL-2003  
DEFINITION Selectively replicating viral vector.  
ACCESSION BD272328.1 GI:33082096  
KEYWORDS UP 2002541761-A/10.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
1 (bases 1 to 56)  
REFERENCE Ramachandra,M. and Shabram,P.W.  
AUTHORS Selectively replicating viral vector  
TITLE Patent: JP 2002541761-A 10 10-DEC-2002;  
JOURNAL CANJI INC

COMMENT  
OS Artificial Sequence  
PN JP 2002541761-A/10  
PD 10-DEC-2002  
PF 14-OCT-1999 JP 2000576027  
PR 15-OCT-1998 US 09/172686  
PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM  
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,C12N5/10,C12N7/00, PC  
C12N15/00,  
PC C12N5/00  
CC complementary 5'-phosphorylated oligonucleotide containing CC  
consensus p53  
CC binding sites(p53CON)  
FH Key Location/Qualifiers  
FT source 1..56  
Location/Qualifiers  
1..56  
/organism="Artificial Sequence".

FEATURES  
Source  
1..56  
/organism="synthetic construct"

ORIGIN /mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

Query Match 100.0%; Score 20; DB 6; Length 56;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 46 GGACATGCCCGGCATGTCC 27

RESULT 21  
LOCUS BV056035  
DEFINITION S212P6623PB2.T0 CZECHII/EI Mus musculus STS genomic, sequence  
ACCESSION BV056035  
VERSION BV056035.1 GI:31171830  
KEYWORDS STS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,  
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.  
TITLE The mosaic structure of variation in the laboratory mouse genome  
JOURNAL Nature 420 (6915), 574-578 (2002)  
MEDLINE 22354684  
PUBMED 12466852

COMMENT Contact: Kerstin Lindblad-Toh  
Whitehead Institute for Biomedical Research, Center for Genome  
Research  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172521477  
Fax: 6172580903  
Email: kersli@genome.wi.mit.edu  
Primer A: None  
Primer B: None  
STS size: 637

Protocol: WGS-discovery: Paired-end low-coverage whole genome shotgun reads  
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS  
reads were placed uniquely on the MGSCV3 C57BL/6J assembly and SNP  
detection was carried out by SSNA-SNP. 225,000 reads were  
annotated  
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J  
and the strain from which the particular read came. The validation  
rate for these SNPs was estimated at approximately 98%.  
Location/Qualifiers

FEATURES  
source 1..637  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="CZECHII/EI"  
/db\_xref="taxon:10090"  
/map="11 23-568 3142702-3143247"  
/clone.lib="CZECHII/EI"  
1..5637

STS  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 11; Length 637;  
Best Local Similarity 95.0%; Pred. No. 3.4e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 497 GGACATGCCCGGCATGTCC 516

RESULT 22

BV056035/c 637 bp DNA linear STS 31-MAY-2003  
LOCUS S212P6623PB2.T0 CZECHII/EI Mus musculus STS genomic, sequence  
DEFINITION tagged site.

ACCESSION BV056035  
VERSION BV056035.1 GI:31171830  
KEYWORDS STS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,  
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.  
TITLE The mosaic structure of variation in the laboratory mouse genome  
JOURNAL Nature 420 (6915), 574-578 (2002)  
MEDLINE 22354684  
PUBMED 12466852

COMMENT Contact: Kerstin Lindblad-Toh  
Whitehead Institute for Biomedical Research, Center for Genome  
Research  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172521477  
Fax: 6172580903  
Email: kersli@genome.wi.mit.edu  
Primer A: None  
Primer B: None  
STS size: 637

Protocol: WGS-discovery: Paired-end low-coverage whole genome shotgun reads  
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS  
reads were placed uniquely on the MGSCV3 C57BL/6J assembly and SNP  
detection was carried out by SSNA-SNP. 225,000 reads were  
annotated  
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J  
and the strain from which the particular read came. The validation  
rate for these SNPs was estimated at approximately 98%.  
Location/Qualifiers

FEATURES  
source 1..637  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="CZECHII/EI"  
/db\_xref="taxon:10090"  
/map="11 23-568 3142702-3143247"  
/clone.lib="CZECHII/EI"  
1..5637

STS  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 11; Length 637;  
Best Local Similarity 95.0%; Pred. No. 3.4e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 516 GGACATGCCCGGCATGTCC 497

RESULT 23  
LOCUS AC021460 146698 bp DNA linear PRI 15-MAR-2002  
DEFINITION Homo sapiens chromosome 15, clone RP11-259F15, complete sequence.  
ACCESSION AC021460  
VERSION AC021460.10 GI:19482349  
KEYWORDS HUG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 15, clone RP11-259F15  
JOURNAL Unpublished

```

REFERENCE
AUTHORS      2 (bases 1 to 146698)
               Birtren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
               Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
               Boguslavsky,L., Bouhgalter,B., Brown,A., Burkett,G., Castle,A.,
               Choebel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
               Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fennetor,J.,
               Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
               Gardina,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
               Howland,J.C., Johnson,R., Jones,C., Kam,L., Karatas,A., Klein,J.,
               Lander,T., Lehoczy,J., Levine,R., Lien,C., Liu,G., Locke,K.,
               MacDonald,P., Margulis,N., McEwan,P., McGuirk,A., McKernan,K.,
               McPheeters,R., Meldrim,J., Menues,L., Morrow,J., Naylor,J.,
               Norman,C.H., O'Connor,T., O'Donnell,P., Olyvar,T.M., Peterson,K.,
               Pierre,N., Plesant,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
               Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
               Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
               Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
               Zimmer,A., and Zody,M.

TITLE
JOURNAL      Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
REFERENCE    Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS      3 (bases 1 to 146698)
               Birtren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
               Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
               Bouhgalter,B., Brown,A., Camarata,J., Campoliano,A., Chang,J.,
               Chazaro,B., Choebel,Y., Colangelo,M., Collins,S., Collamore,A.,
               Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
               Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardina,S.,
               Glinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C.,
               Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
               Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
               Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
               Maclean,C., MacDonald,P., Major,J., Margie,N., Matthews,C.,
               McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menues,L.,
               Milnova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
               Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
               Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Pollara,V.,
               Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
               Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
               Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
               Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
               Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
               Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
               Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL      Direct Submission
COMMENT      Submitted (15-MAR-2002) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA
               On Mar 15, 2002 this sequence version replaced gi:17223355.
               All repeats were identified using RepeatMasker:
               Smit, A.F.A. & Green, P. (1996-1997)
               http://ftp.genome.washington.edu/RM/RepeatMasker.html

               Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Center code: WIBR
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence.submissions@genome.wi.mit.edu

               Project Information
               Center project name: L5195
               Center clone name: 259_F_15

FEATURES
SOURCE
               Location/Qualifiers
               1. 146698
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="15"
               /map="15"
               /clone="RP11-259P15"
               /clone.lib="RP11-11 Human Male BAC"
               repeat_region
               59..119
               /rpt_family="MIR"
               repeat_region
               216..561
               /rpt_family="AluDb"
               repeat_region
               complement(562..732)
               /rpt_family="L2"
               repeat_region
               1082..1119
               /rpt_family="AT-rich"
               repeat_region
               1120..1415
               /rpt_family="AluDb"
               complement(1417..1696)
               /rpt_family="L1MB8"
               repeat_region
               complement(1697..1985)
               /rpt_family="AluSc"
               repeat_region
               complement(1986..2138)
               /rpt_family="L1MB8"
               repeat_region
               2798..2881
               /rpt_family="MIR"
               repeat_region
               2958..3108
               /rpt_family="LTR33"
               complement(3113..3376)
               /rpt_family="L1PBA"
               repeat_region
               3397..3639
               /rpt_family="LTR33A"
               complement(3640..3977)
               /rpt_family="THE1B"
               repeat_region
               3978..4097
               /rpt_family="LTR33A"
               complement(4257..4566)
               /rpt_family="AluSc"
               repeat_region
               complement(4971..5131)
               /rpt_family="FRAM"
               repeat_region
               5255..5494
               /rpt_family="AluDo"
               repeat_region
               5649..5963
               /rpt_family="AluDo"
               repeat_region
               6659..6791
               /rpt_family="AluSc"
               repeat_region
               8068..8244
               /rpt_family="AluDo"
               repeat_region
               9323..9497
               /rpt_family="AluDo"
               repeat_region
               9600..9871
               /rpt_family="L3"
               repeat_region
               9875..10038
               /rpt_family="L3"
               repeat_region
               10045..10340
               /rpt_family="AluDo"
               repeat_region
               10371..10419
               /rpt_family="AluSc"
               repeat_region
               complement(10479..10844)
               /rpt_family="AT-rich"
               repeat_region
               complement(11081..11323)
               /rpt_family="L2"
               repeat_region
               complement(11750..12055)
               /rpt_family="L1MC4a"
               repeat_region
               /rpt_family="AluSc"
               repeat_region
               12281..12405
               /rpt_family="MIR"
               repeat_region
               12691..12764
               /rpt_family="MIR"
               repeat_region
               13182..13457
               /rpt_family="TATATG)n"
               repeat_region
               13461..13484
               /rpt_family="AluSc"
               repeat_region
               13568..13651
               /rpt_family="CMA)n"
               repeat_region
               /rpt_family="MIR3"
               complement(13859..14002)
               /rpt_family="MIR3"
               complement(14683..14795)
               /rpt_family="MIR"
               repeat_region
               14892..15185
               /rpt_family="AluSc"
               repeat_region
               complement(16265..16367)
               /rpt_family="MIR"
               repeat_region
               16368..16405
               /rpt_family="TG)n"
               repeat_region
               complement(16406..16541)

```





```

repeat_region /rpt_family="L1MB8"
complement(1697. .1985)
repeat_region /rpt_family="AluSc"
complement(1986. .2138)
repeat_region /rpt_family="L1MB8"
2798. .2881
repeat_region /rpt_family="MIR"
2958. .3108
repeat_region /rpt_family="LTR33"
complement(3113. .3376)
repeat_region /rpt_family="L1PBA"
3397. .3639
repeat_region /rpt_family="LTR33A"
complement(3640. .3977)
repeat_region /rpt_family="THE1B"
3978. .4097
repeat_region /rpt_family="LTR33A"
complement(4257. .4566)
repeat_region /rpt_family="AluSc"
complement(4971. .5131)
repeat_region /rpt_family="FRAM"
5255. .5494
repeat_region /rpt_family="AluUo"
5649. .5963
repeat_region /rpt_family="AluSc"
6659. .6791
repeat_region /rpt_family="AluUo"
8068. .8244
repeat_region /rpt_family="AluUo"
9323. .9497
repeat_region /rpt_family="L3"
9600. .9871
repeat_region /rpt_family="L3"
9875. .10038
repeat_region /rpt_family="AluUo"
10045. .10340
repeat_region /rpt_family="AluSc"
10371. .10419
repeat_region /rpt_family="AT rich"
complement(10475. .10844)
repeat_region /rpt_family="L2"
complement(11081. .11323)
repeat_region /rpt_family="L1MC4a"
complement(11750. .12055)
repeat_region /rpt_family="AluSc"
12281. .12405
repeat_region /rpt_family="MIR"
12691. .12764
repeat_region /rpt_family=" (TATATG)n"
13182. .13457
repeat_region /rpt_family="AluSc"
13461. .13484
repeat_region /rpt_family=" (CAA)n"
13568. .13651
repeat_region /rpt_family="MIR3"
complement(13859. .14002)
repeat_region /rpt_family="MIR3"
complement(14683. .14795)
repeat_region /rpt_family="MIR"
14892. .15185
repeat_region /rpt_family="AluSc"
complement(16265. .16367)
repeat_region /rpt_family="MIR"
16368. .16405
repeat_region /rpt_family=" (TG)n"
complement(16406. .16541)
repeat_region /rpt_family="MIR"
complement(16553. .16824)
repeat_region /rpt_family="L1MC4a"
complement(16974. .17273)
repeat_region /rpt_family="AluSc"
complement(17892. .18185)
repeat_region /rpt_family="AluSc"

```

```

repeat_region complement(19150. .19457)
repeat_region /rpt_family="AluU"
complement(19464. .19526)
repeat_region /rpt_family="MER5A"
19527. .19922
repeat_region /rpt_family="L1MA9"
19947. .20241
repeat_region /rpt_family="AluUo"
20267. .20309
repeat_region /rpt_family="Charlie1"
20321. .20702
repeat_region /rpt_family="L1MD2"
20713. .20824
repeat_region /rpt_family="MER5A"
complement(21299. .21499)
repeat_region /rpt_family="MIR"
21650. .21763
repeat_region /rpt_family="MIR"
21875. .21909
repeat_region /rpt_family="MER5B"
complement(21910. .22226)
repeat_region /rpt_family="AluSc"
22227. .22484
repeat_region /rpt_family="MER5B"
22757. .23190
repeat_region /rpt_family="MLT1J"
complement(23470. .23775)
repeat_region /rpt_family="AluSc"
23781. .23803
repeat_region /rpt_family=" (TTTAA)n"
24191. .24257
repeat_region /rpt_family=" (TG)n"
25248. .25279
repeat_region /rpt_family="T-rich"

```

Query Match 92.0%; Score 18.4; DB 9; Length 146698;  
 Best Local Similarity 95.0%; Pred.No. 2.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGCATGTCC 20  
 Db 106527 GGACATGCTCGCATGTCC 106508

RESULT 25  
 AL359638 159478 bp DNA linear HTG 10-JUL-2001  
 LOCUS Homo sapiens chromosome 6 clone RP11-769A14, 14 unordered pieces.  
 DEFINITION  
 ACCESSION AL359638  
 VERSION HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Sims,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 COMMENT On Jan 26, 2001 this sequence version replaced gi:9864217.  
 ----- Genome Center  
 Center: Sanger Centre  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 ----- Project Information  
 Center project name: BA769A14  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; 108752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 153705 bases at least Q40  
 Consensus quality: 155925 bases at least Q30  
 Consensus quality: 157184 bases at least Q20  
 Insert size: 158178; sum-of-contigs  
 Insert size: 154847; agarose-fp  
 Quality coverage: 4.26x in Q20 bases; sum-of-contigs Quality  
 coverage: 4.51x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 14 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1      6892: contig of 6892 bp in length
*      6893      6992: gap of 100 bp
*      6993      14369: contig of 7377 bp in length
*      14370      14469: gap of 100 bp
*      14469      19375: contig of 4806 bp in length
*      19375      19475: gap of 100 bp
*      19475      33036: contig of 13561 bp in length
*      33036      33137: gap of 100 bp
*      33137      42023: contig of 8887 bp in length
*      42023      42123: gap of 100 bp
*      42123      73902: contig of 31779 bp in length
*      73902      74002: gap of 100 bp
*      74002      84868: contig of 10866 bp in length
*      84868      84968: gap of 100 bp
*      84968      11588: contig of 26620 bp in length
*      11588      11688: gap of 100 bp
*      11688      131172: contig of 19484 bp in length
*      131172      131272: gap of 100 bp
*      131272      134785: contig of 3513 bp in length
*      134785      134786: gap of 100 bp
*      134786      138106: contig of 3221 bp in length
*      138106      138207: gap of 100 bp
*      138207      147740: contig of 9534 bp in length
*      147740      147841: gap of 100 bp
*      147841      151669: contig of 3329 bp in length
*      151669      151769: gap of 100 bp
*      151769      159478: contig of 7709 bp in length.
*      159478      159478: contig of 7709 bp in length.
  
```

## FEATURES

```

source
1..159478
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="6"
  /clone="RP11-769A14"
  /clone_1fb="RPC1-11.3"
  1..6892
    /note="assembly_fragment:00602"
    fragment_chain:1
  6893..14369
    /note="assembly_fragment:02031"
    fragment_chain:1
  14470..19375
    /note="assembly_fragment:01886"
    fragment_chain:1
  19476..33036
    /note="assembly_fragment:01758"
    fragment_chain:1
  33137..42023
    /note="assembly_fragment:01839"
    fragment_chain:1
  42124..73902
    /note="assembly_fragment:00383"
    fragment_chain:2
  74003..84868
    /note="assembly_fragment:01701"
    fragment_chain:2
  84969..11588
  
```

```

/note="assembly_fragment:00704"
fragment_chain:3
111689..131172
/note="assembly_fragment:01843"
fragment_chain:3
131273..134785
/note="assembly_fragment:00807"
fragment_chain:4
134886..138106
/note="assembly_fragment:01302"
fragment_chain:4
138207..147740
/note="assembly_fragment:01109"
fragment_chain:5
147841..151669
/note="assembly_fragment:01940"
fragment_chain:5
151770..159478
/note="assembly_fragment:00573"
clone_end:17
vector_side:right"
  
```

## ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 159478;  
 Best Local Similarity 95.0%; Pred. No. 2.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Cy      1 GGACATGCCGGGCATGTC 20
Db      139362 GGACATGCTCGGCATGTC 139361
  
```

## RESULT 26

```

AL359638/c      159478 bp      DNA      linear      HTG 10-JUL-2001
LOCUS      Homo sapiens chromosome 6 clone RP11-769A14, 14 unordered pieces.
DEFINITION
ACCESSION      AL359638
VERSION      AL359638.7 GI:12539703
KEYWORDS
SOURCE      HTG; HTGS_PHASE1; HTGS_CANCELLED.
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  
```

## REFERENCE

```

AUTHORS      Sims, S.
TITLE      Direct Submission
JOURNAL      Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            requests: clonerequest@sanger.ac.uk
            On Jan 26, 2001 this sequence version replaced gi:9864217.
COMMENT
            ----- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ----- Project Information
            Center project name: ba769A14
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; 108752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 153705 bases at least Q40
            Consensus quality: 155925 bases at least Q30
            Consensus quality: 157184 bases at least Q20
            Insert size: 158178; sum-of-contigs
            Insert size: 154847; agarose-fp
            Quality coverage: 4.26x in Q20 bases; sum-of-contigs Quality
            coverage: 4.51x in Q20 bases; agarose-fp
  
```

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 14 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 6892: contig of 6892 bp in length
6893 6992: gap of 100 bp
6993 14369: contig of 7377 bp in length
14370 14469: gap of 100 bp
14470 19375: contig of 4906 bp in length
19376 19475: gap of 100 bp
19476 33036: contig of 13561 bp in length
33037 33136: gap of 100 bp
33137 42023: contig of 8887 bp in length
42024 42123: gap of 100 bp
42124 73902: contig of 31779 bp in length
73903 74002: gap of 100 bp
74003 84868: contig of 10866 bp in length
84869 84969: gap of 100 bp
84969 111588: contig of 26620 bp in length
111589 111689: gap of 100 bp
111689 131172: contig of 19484 bp in length
131173 131272: gap of 100 bp
131273 134785: contig of 3513 bp in length
134786 134885: gap of 100 bp
134886 138106: contig of 3221 bp in length
138107 138206: gap of 100 bp
138207 147740: contig of 9534 bp in length
147741 147840: gap of 100 bp
147841 151693: contig of 3829 bp in length
151694 151770: gap of 100 bp
151770 159478: contig of 7709 bp in length.
Location/Qualifiers
1. 159478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone_id="RPC1-11.3"
1. 6852
/note="assembly_fragment:00602
fragment_chain:1"
6993. 14369
/note="assembly_fragment:02031
fragment_chain:1"
14470. 19375
/note="assembly_fragment:01886
fragment_chain:1"
19476. 33036
/note="assembly_fragment:01758
fragment_chain:1"
33137. 42023
/note="assembly_fragment:01839
fragment_chain:1"
42124. 73902
/note="assembly_fragment:00383
fragment_chain:2"
74003. 84868
/note="assembly_fragment:01701
fragment_chain:2"
84969. 111588
/note="assembly_fragment:00704
fragment_chain:3"
111689. 131172
/note="assembly_fragment:01843
fragment_chain:3"
131273. 134785
/note="assembly_fragment:00807
fragment_chain:4"
134886. 138106
/note="assembly_fragment:01302
fragment_chain:4"
138207. 147740

```

```

FEATURES
source
1. 159478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone_id="RPC1-11.3"
1. 6852
/note="assembly_fragment:00602
fragment_chain:1"
6993. 14369
/note="assembly_fragment:02031
fragment_chain:1"
14470. 19375
/note="assembly_fragment:01886
fragment_chain:1"
19476. 33036
/note="assembly_fragment:01758
fragment_chain:1"
33137. 42023
/note="assembly_fragment:01839
fragment_chain:1"
42124. 73902
/note="assembly_fragment:00383
fragment_chain:2"
74003. 84868
/note="assembly_fragment:01701
fragment_chain:2"
84969. 111588
/note="assembly_fragment:00704
fragment_chain:3"
111689. 131172
/note="assembly_fragment:01843
fragment_chain:3"
131273. 134785
/note="assembly_fragment:00807
fragment_chain:4"
134886. 138106
/note="assembly_fragment:01302
fragment_chain:4"
138207. 147740

```

```

misc_feature
147841..151669
/note="assembly_fragment:01940
fragment_chain:5"
151770..159478
/note="assembly_fragment:00573
clone_end:17
vector_side:right"
ORIGIN

```

```

Query Match 92.0%; Score 18.4; DB 2; Length 159478;
Best Local Similarity 95.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OR
1 GGACATGCCCGGCGATGTCC 20
|||||
Db 139381 GGACATGCCCGGCGATGTCC 139362

```

#### RESULT 27

```

LOCUS HS44A20 165090 bp DNA linear PRI 04-MAR-2003
DEFINITION Human DNA sequence from clone RPI-44A20 on chromosome 6q23.1-24.3,
complete sequence.
ACCESSION AL035086
VERSION AL035086.12 GI:4741478
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 165090)
Matthews, L.
Direct Submission
Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humany@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 4, 1999 this sequence version replaced gi:4680389.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humany@sanger.ac.uk
-----

```

#### COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>

RPI-44A20 is from the library RPC1-1 constructed by the group of Pieter de Jong. For further details see <http://www.choyl.org/bacpac/home.htm>

VECTOR: pCYPAC2.

#### FEATURES

Location/Qualifiers

source 1. 165090

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="q23.1-24.3"  
/clone="RP1-44A20"  
/clone\_lib="RP1-1"

## ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 165090;  
Best Local Similarity 95.0%; Pred. No. 2.3e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20

Db 67049 GGACATGCCCGGCATGTCC 67068

## RESULT 28

HS44A20

LOCUS 165090 bp DNA linear PRI 04-MAR-2003  
DEFINITION Human DNA sequence from clone RP1-44A20 on chromosome 6q23.1-24.3,  
complete sequence.  
ACCESSION AL035086  
VERSION AL035086  
KEYWORDS HTG. GI:4741478  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Matthews, L.  
1 (bases 1 to 165090)

## REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On May 4, 1999 this sequence version replaced gi:4680389.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30);  
an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6

RP1-44A20 is from the library RP1-1 constructed by the group of  
Pietter de Jong. For further details see  
http://www.choi.org/bacpac/home.htm  
VECTOR: pCYPAC2

FEATURES  
Source  
Location/Qualifiers  
1. 165090

/organism="Homo sapiens"

/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="q23.1-24.3"  
/clone="RP1-44A20"  
/clone\_lib="RP1-1"

## ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 165090;  
Best Local Similarity 95.0%; Pred. No. 2.3e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20

Db 67068 GGACATGCCCGGCATGTCC 67049

## RESULT 29

AC126323

LOCUS

DEFINITION Homo sapiens chromosome 15, clone RP11-299H22, complete sequence.  
ACCESSION AC126323  
VERSION AC126323.6 GI:24850482  
KEYWORDS HTG. GI:24850482  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 170132)  
Homo sapiens chromosome 15, clone RP11-299H22  
Unpublished  
2 (bases 1 to 170132)

## REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

## COMMENT

Submitted (05-JUN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 170132)  
Direct Submission  
Britten, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barn, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindsad-Roh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunhahang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Tefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zemdek, L., Zimmer, A. and Zody, W.

## TITLE

JOURNAL

REFERENCE

AUTHORS

Britten, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barn, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindsad-Roh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunhahang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Tefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zemdek, L., Zimmer, A. and Zody, W.

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

Wyman, D., Young, G., Zainoun, J., Zemбек, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (31-OCT-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 170132)  
 Birren, B., Nisbaum, C., Lander, B., Ali, A., Allen, N., Anderson, S.,  
 Barna, N., Baetsen, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,  
 Chazara, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
 Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galegan, J.,  
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
 Hagos, B., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,  
 Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,  
 Matthews, C., McCarthy, M., Meldrum, J., Meneses, L., Mihova, T.,  
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Notbu, C.,  
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
 Rife, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,  
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, K.,  
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zemбек, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (11-NOV-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Nov 11, 2002 this sequence version replaced gi:24431891.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L27748  
 Center clone name: 299\_H\_22

**FEATURES**  
**SOURCE**

----- Location/Qualifiers  
 1. 170132  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="15"  
 /map="15"  
 /clone="RP11-299H22"  
 /clone\_id="RPC1-11 Human Male BAC"  
 complement(92..287)  
 repeat\_region  
 /rpt\_family="MIR"  
 1956..2179  
 /rpt\_family="MIR"  
 complement(3303..3533)  
 repeat\_region  
 /rpt\_family="MIR"  
 5010..5166  
 /rpt\_family="MIR"  
 complement(5503..5759)  
 repeat\_region  
 /rpt\_family="L2"  
 6342..6441  
 /rpt\_family="L1MD3"  
 7226..7270  
 /rpt\_family="poly(pyrimidine)"  
 complement(7275..7354)  
 repeat\_region  
 /rpt\_family="LTR16A"  
 complement(7491..7674)  
 repeat\_region  
 /rpt\_family="L1MCS5"  
 7675..7710  
 /rpt\_family="TTTA)n"  
 complement(7711..7782)  
 repeat\_region  
 /rpt\_family="L1MCS5"  
 complement(7784..7905)  
 repeat\_region  
 /rpt\_family="L1MA5A"  
 8239..9936  
 repeat\_region

repeat\_region  
 9937..10293  
 /rpt\_family="L1MA4"  
 repeat\_region  
 10295..10897  
 /rpt\_family="MSTC"  
 repeat\_region  
 10895..10960  
 /rpt\_family="L1MA4"  
 complement(10961..11197)  
 repeat\_region  
 /rpt\_family="L1PB1"  
 11198..12677  
 /rpt\_family="L1PB1"  
 12733..12778  
 /rpt\_family="TA)n"  
 complement(12917..12968)  
 repeat\_region  
 /rpt\_family="L3"  
 13077..13199  
 /rpt\_family="MIR3"  
 complement(13593..13747)  
 repeat\_region  
 /rpt\_family="MIR"  
 14902..14923  
 /rpt\_family="GC\_rich"  
 15271..15404  
 /rpt\_family="GC\_rich"  
 16661..16696  
 /rpt\_family="U2"  
 complement(16882..17005)  
 repeat\_region  
 /rpt\_family="MIR"  
 17405..17491  
 /rpt\_family="L2"  
 complement(17561..17635)  
 repeat\_region  
 /rpt\_family="MIR"  
 17712..17760  
 /rpt\_family="CT\_rich"  
 complement(17975..18252)  
 repeat\_region  
 /rpt\_family="MER2"  
 complement(18688..18823)  
 repeat\_region  
 /rpt\_family="MIR"  
 complement(18971..19199)  
 repeat\_region  
 /rpt\_family="MIR"  
 complement(19842..20085)  
 repeat\_region  
 /rpt\_family="L1ME3A"  
 20168..20449  
 /rpt\_family="MLT1A1"  
 complement(20451..20627)  
 repeat\_region  
 /rpt\_family="AluDb"  
 20661..20795  
 /rpt\_family="MLT1B"  
 complement(20804..20907)  
 repeat\_region  
 /rpt\_family="L1ME3A"  
 20975..21343  
 /rpt\_family="MER47A"  
 21413..21523  
 /rpt\_family="L3"  
 21864..21953  
 /rpt\_family="MIR"  
 complement(21977..22294)  
 repeat\_region  
 /rpt\_family="AluSg"  
 22897..23037  
 /rpt\_family="MIR3"  
 complement(23372..23830)  
 repeat\_region  
 /rpt\_family="MLT1C"  
 complement(24687..25005)  
 repeat\_region  
 /rpt\_family="LTR16A"  
 25827..26077  
 /rpt\_family="MER33"  
 26921..26952  
 /rpt\_family="AT\_rich"  
 27182..27237  
 /rpt\_family="MIR"  
 complement(27257..27319)  
 repeat\_region  
 /rpt\_family="MIR"  
 27889..27923  
 repeat\_region  
 /rpt\_family="AT\_rich"



repeat\_region complement(10961. .11197)  
 /rpt\_family="L1PB1"  
 repeat\_region 11198. .12677  
 /rpt\_family="L1PB1"  
 repeat\_region 12733. .12778  
 /rpt\_family="(TA)n"  
 repeat\_region complement(12917. .12968)  
 /rpt\_family="L3"  
 repeat\_region 13077. .13199  
 /rpt\_family="MIR3"  
 repeat\_region complement(13593. .13747)  
 /rpt\_family="MIR"  
 repeat\_region 14902. .14923  
 /rpt\_family="GC\_rich"  
 repeat\_region 15271. .15404  
 /rpt\_family="GC\_rich"  
 repeat\_region 16661. .16696  
 /rpt\_family="U2"  
 repeat\_region complement(16882. .17005)  
 /rpt\_family="MIR"  
 repeat\_region 17405. .17491  
 /rpt\_family="L2"  
 repeat\_region complement(17561. .17635)  
 /rpt\_family="MIR"  
 repeat\_region 17712. .17760  
 /rpt\_family="CT\_rich"  
 repeat\_region complement(17975. .18252)  
 /rpt\_family="MER2"  
 repeat\_region complement(18688. .18823)  
 /rpt\_family="MIR"  
 repeat\_region complement(18971. .19199)  
 /rpt\_family="MIR"  
 repeat\_region complement(19842. .20085)  
 /rpt\_family="L1M3A"  
 repeat\_region 20168. .20449  
 /rpt\_family="MLT1A1"  
 repeat\_region complement(20451. .20627)  
 /rpt\_family="AluDb"  
 repeat\_region 20661. .20795  
 /rpt\_family="MLT1B"  
 repeat\_region complement(20804. .20907)  
 /rpt\_family="L1M3A"  
 repeat\_region 20979. .21343  
 /rpt\_family="MER47A"  
 repeat\_region 21413. .21523  
 /rpt\_family="L3"  
 repeat\_region 21864. .21953  
 /rpt\_family="MIR"  
 repeat\_region complement(21977. .22294)  
 /rpt\_family="AluSg"  
 repeat\_region 22897. .23037  
 /rpt\_family="MIR3"  
 repeat\_region complement(23372. .23830)  
 /rpt\_family="MLT1C"  
 repeat\_region complement(24687. .25005)  
 /rpt\_family="LTR16A"  
 repeat\_region 25827. .26077  
 /rpt\_family="MER33"  
 repeat\_region 26921. .26952  
 /rpt\_family="AT\_rich"  
 repeat\_region 27182. .27237  
 /rpt\_family="MIR"  
 repeat\_region complement(27257. .27319)  
 /rpt\_family="MIR"  
 repeat\_region 27889. .27923  
 /rpt\_family="AT\_rich"  
 repeat\_region 28176. .28351

Query March 92.0% Score 18.4; DB 9; Length 170132;  
 Best Local Similarity 95.0%; Pred.No.2.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGACATGCCCGGCGCATGTCC 20

Db 113566 GGACATGCCCGGCGCATGTCC 113547  
 RESULT 31  
 AL670675 173987 bp DNA linear ROD 12-JUL-2002  
 LOCUS Mouse DNA sequence from clone RP23-263N18 on chromosome X, complete  
 DEFINITION sequence.  
 AL670675  
 ACCESSION AL670675.7 GI:21531153  
 VERSION HTG.  
 KEYWORDS Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 173987)  
 REFERENCE Wray,P.  
 AUTHORS Direct Submission  
 TITLE Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
 JOUENAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
 On Jun 21, 2002 this sequence version replaced gi:20330273.  
 COMMENT  
 Center: Wellcome Trust Sanger Institute  
 Genome Center  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; Sw.,  
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-263N18 is  
 from the RPCI-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6.  
 FEATURES  
 Location/Qualifiers  
 1. 173987  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP23-263N18"  
 /clone\_id="RPCI-23"  
 ORIGIN  
 Query March 92.0% Score 18.4; DB 10; Length 173987;  
 Best Local Similarity 95.0%; Pred.No.2.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 95248 GGACATGCCCGGCGCATGTCC 95267  
 QY 1 GGACATGCCCGGCGCATGTCC 20  
 AL670675 173987 bp DNA linear ROD 12-JUL-2002  
 LOCUS Mouse DNA sequence from clone RP23-263N18 on chromosome X, complete  
 DEFINITION sequence.  
 AL670675  
 ACCESSION AL670675.7 GI:21531153  
 VERSION HTG.  
 KEYWORDS Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 173987)  
 REFERENCE Wray,P.  
 AUTHORS Direct Submission  
 TITLE Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
 JOUENAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
 On Jun 21, 2002 this sequence version replaced gi:20330273.  
 COMMENT  
 Center: Wellcome Trust Sanger Institute  
 Genome Center  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; Sw.,  
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-263N18 is  
 from the RPCI-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6.  
 FEATURES  
 Location/Qualifiers  
 1. 173987  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP23-263N18"  
 /clone\_id="RPCI-23"  
 ORIGIN  
 Query March 92.0% Score 18.4; DB 10; Length 173987;  
 Best Local Similarity 95.0%; Pred.No.2.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DEFINITION Mouse DNA sequence from clone RP23-263N18 on chromosome X, complete sequence.  
 ACCESSION AL670675  
 VERSION AL670675.7 GI:21531153  
 KEYWORDS HTG:  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 173987)  
 REFERENCE Wray, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 On Jun 21, 2002 this sequence version replaced gi:20330273.  
 COMMENT ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 from the RP23-263N18 Mouse PAC library constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6.  
 FEATURES  
 source location/Qualifiers  
 1..173987  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP23-263N18"  
 /clone\_11b="RP23-263N18"  
 /clone\_11b="RP23-263N18"  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 10; Length 173987;  
 Best Local Similarity 95.0%; Pred. No. 2.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCATGTC 20  
 Db 95267 GGACATGCCCGGCATGTC 95248  
 RESULT 33  
 AL773522/c  
 LOCUS Mouse DNA sequence from clone RP23-148P20 on chromosome 4, complete sequence.  
 DEFINITION AL773522 185592 bp DNA linear ROD 25-FEB-2003  
 ACCESSION AL773522  
 VERSION AL773522.15 GI:28564418  
 KEYWORDS HTG:  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 185592)  
 AUTHORS Philimore, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-FEB-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 On Feb 25, 2003 this sequence version replaced gi:28268662.  
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
 COMMENT ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 from the RP23-148P20 Mouse BAC library constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6.  
 FEATURES  
 source location/Qualifiers  
 1..185592  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-148P20"  
 /clone\_11b="RP23-148P20"  
 /clone\_11b="RP23-148P20"  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 10; Length 185592;  
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCATGTC 20  
 Db 140068 GGACATGCCCGGCATGTC 140067  
 RESULT 34  
 AL773522/c  
 LOCUS Mouse DNA sequence from clone RP23-148P20 on chromosome 4, complete sequence.  
 DEFINITION AL773522 185592 bp DNA linear ROD 25-FEB-2003  
 ACCESSION AL773522  
 VERSION AL773522.15 GI:28564418  
 KEYWORDS HTG:  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.



REFERENCE 1 (bases 1 to 185592)  
 AUTHORS Phillimore, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-FEB-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 COMMENT On Feb 25, 2003 this sequence version replaced gi:28268662. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
 -----  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 -----  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, SW, SWISSPROT, Tr, TREMBL, Wp: WORMPEP, Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-148P20 is from the RPCI-23 Mouse BAC Library.  
 constructed by the group of Pieter de Jong.  
 For further details see http://www.choiri.org/bacpac/home.htm  
 VECTOR: pBAC3.6.  
 Location/Qualifiers  
 1. 185592  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-148P20"  
 /clone\_11b="RPCI-23"  
 /clone\_11b="RPCI-23"  
 -----  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 10; Length 185592;  
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 -----  
 Oy 1 GGACATGCCCGGCGATGCC 20  
 |||  
 Db 140087 GGGCATGCCCGGCGATGCC 140068  
 -----  
 RESULT 35  
 AC139327  
 LOCUS Mus musculus chromosome UNK clone RP24-362N3, WORKING DRAFT  
 DEFINITION AC139327 191128 bp DNA linear HTG 21-FEB-2003  
 SEQUENCE 5 unordered pieces.  
 ACCESSION AC139327  
 VERSION AC139327.2 GI:28173253  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 191128)  
 McPherson, J.D. and Waterston, R.H.

TITLE The sequence of Mus musculus clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 191128)  
 AUTHORS McPherson, J.D. and Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JUN-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 3 (bases 1 to 191128)  
 AUTHORS McPherson, J.D. and Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-FEB-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 COMMENT On Jan 31, 2003 this sequence version replaced gi:28144456.  
 -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUSGC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 Contact: submissions@watson.wustl.edu  
 Project Information  
 Center project name: M\_BB0362N03  
 -----  
 Summary Statistics  
 Sequencing vector: M13; 0%  
 Sequencing vector: plasmid; 100%  
 Chemistry: Dye-primer ET; 0% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 190351 bases at least Q40  
 Consensus quality: 190551 bases at least Q30  
 Consensus quality: 190792 bases at least Q20  
 Insert size: 164000; agarose-fp  
 Insert size: 193202; sum-of-contigs  
 Quality coverage: 13.72 in Q20 bases; agarose-fp  
 Quality coverage: 12.92 in Q20 bases; sum-of-contigs  
 -----  
 NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
 1 8971: contig of 8971 bp in length  
 \* 8972 9071: gap of unknown length  
 \* 9072 26264: contig of 17193 bp in length  
 \* 26265 26364: gap of unknown length  
 \* 26365 59759: contig of 33395 bp in length  
 \* 59760 59859: gap of unknown length  
 \* 59860 101845: contig of 41986 bp in length  
 \* 101846 101945: gap of unknown length  
 \* 101946 191128: contig of 89183 bp in length.  
 Location/Qualifiers  
 1. 191128  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="UNK"  
 /clone="RP24-362N3"  
 1. 8971  
 /note="assembly\_name:Contig47"  
 misc\_feature 9072..26264  
 /note="assembly\_name:Contig48"  
 misc\_feature 26365..59759  
 /note="assembly\_name:Contig49"  
 misc\_feature 59860..101845  
 /note="assembly\_name:Contig50"  
 misc\_feature 101946..191128  
 /note="assembly\_name:Contig51"  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 2; Length 191128;

Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCATGTC 20

Db 99262 GGACCTGCCCGGCATGTC 99281

RESULT 36  
AC139327/c

LOCUS AC139327 191128 bp DNA linear HTG 21-FEB-2003  
DEFINITION Mus musculus chromosome UNK clone RP24-362N3, WORKING DRAFT  
SEQUENCE, 5 unordered pieces.

ACCESSION AC139327

VERSION AC139327.2 GI:28173253

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 191128)

McPherson, J.D. and Waterston, R.H.

The sequence of Mus musculus clone

2 (bases 1 to 191128)

McPherson, J.D. and Waterston, R.H.

Submitted (30-JUN-2003) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 191128)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (21-FEB-2003) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: [submissions@waterston.wustl.edu](mailto:submissions@waterston.wustl.edu)

----- Project Information -----

Center project name: M\_BB0362M03

----- Summary Statistics -----

Sequencing vector: M13; 0%

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 190351 bases at least Q40

Consensus quality: 190555 bases at least Q30

Consensus quality: 190792 bases at least Q20

Insert size: 16400; agarose-fp

Insert size: 19320; sum-of-contigs

Quality coverage: 13.72 in Q20 bases; agarose-fp

Quality coverage: 12.92 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently

consists of 5 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

1 8971: contig of 8971 bp in length

\* 8972 9071: gap of unknown length

\* 9072 26264: contig of 17193 bp in length

\* 26265 26364: gap of unknown length

\* 26365 59759: contig of 33395 bp in length

\* 59760 59859: gap of unknown length

\* 59860 101845: contig of 41986 bp in length

\* 101846 101945: gap of unknown length  
\* 101946 191128: contig of 89183 bp in length.  
Location/Qualifiers  
1. 191128

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="XNK"

/clone="RP24-362N3"

1. 8971

/note="assembly\_name:Contig47"

9072. 26264

/note="assembly\_name:Contig48"

26365. 59759

/note="assembly\_name:Contig49"

59860. 101845

/note="assembly\_name:Contig50"

101946. 191128

/note="assembly\_name:Contig51"

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

\* NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1412: contig of 1412 bp in length
1413 1512: gap of unknown length
1513 3635: contig of 2123 bp in length
3636 3735: gap of unknown length
3736 4999: contig of 1264 bp in length
5000 5099: gap of unknown length
5100 7667: contig of 2568 bp in length
7668 7767: gap of unknown length
7768 10116: contig of 2349 bp in length
10117 10216: gap of unknown length
10217 14002: contig of 3786 bp in length
14003 14102: gap of unknown length
14103 17164: contig of 3062 bp in length
17165 17264: gap of unknown length
17265 20585: contig of 3321 bp in length
20586 20685: gap of unknown length
20686 24370: contig of 3685 bp in length
24371 24470: gap of unknown length
24471 29951: contig of 5481 bp in length
29952 30051: gap of unknown length
30052 37837: contig of 7786 bp in length
37838 37937: gap of unknown length
37938 48151: contig of 10214 bp in length
48152 48251: gap of unknown length
48252 61495: contig of 13244 bp in length
61496 61595: gap of unknown length
61596 73444: contig of 11749 bp in length
73445 73444: gap of unknown length
73445 85338: contig of 11894 bp in length
85339 85439: gap of unknown length
85439 97336: contig of 11898 bp in length
97337 97436: gap of unknown length
97437 109646: contig of 12210 bp in length
109647 109747: gap of unknown length
109747 123544: contig of 13798 bp in length
123545 123644: gap of unknown length
123645 139857: contig of 16213 bp in length
139858 139957: gap of unknown length
139958 154527: contig of 14570 bp in length
154528 154627: gap of unknown length
154628 174231: contig of 19604 bp in length
174232 174331: gap of unknown length
174332 202227: contig of 27896 bp in length.
174332

```

## FEATURES

```

source
1. .202227
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/clone="RP11-546114"
1. .1412
/clone="assembly_name:Contig6"
1513. .3635
/clone="assembly_name:Contig7"
3736. .4999
/clone="assembly_name:Contig8"
5100. .7667
/clone="assembly_name:Contig9"
7768. .10116
/clone="assembly_name:Contig10"
10217. .14002
/clone="assembly_name:Contig11"
14103. .17164
/clone="assembly_name:Contig12

```

```

clone end:17
vector_side:left"
misc_feature
17265..20585
/clone="assembly_name:Contig13"
misc_feature
20686..24370
/clone="assembly_name:Contig14"
misc_feature
24471..29951
/clone="assembly_name:Contig15"
misc_feature
30052..37837
/clone="assembly_name:Contig16"
misc_feature
37938..48151
/clone="assembly_name:Contig17"
misc_feature
48252..61495
/clone="assembly_name:Contig18"
misc_feature
61596..73344
/clone="assembly_name:Contig19"
misc_feature
73445..85338
/clone="assembly_name:Contig20
clone end:SP6
vector_side:right"
85439..97336
/clone="assembly_name:Contig21"
misc_feature
97437..109646
/clone="assembly_name:Contig22"
misc_feature
109747..123544
/clone="assembly_name:Contig23"
misc_feature
123645..139857
/clone="assembly_name:Contig24"
misc_feature
139958..154527
/clone="assembly_name:Contig25"
misc_feature
154628..174231
/clone="assembly_name:Contig26"
174332..202227
/clone="assembly_name:Contig27"

```

## ORIGIN

```

Query Match 92.0%; Score 18.4; DB 2; Length 202227;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 GGACATGCCCGGCGATGTC 20
|||||

```

```

Db 184949 GGACATGCCCTGGCGATGCC 184968
|||||

```

## RESULT 38

```

AC068126/c
LOCUS

```

```

DEFINITION AC068126 202227 bp DNA linear HTG 07-JUL-2000
SEQUENCE, 22 unordered pieces.
ACCESSION AC068126
VERSION AC068126.2 GI:7705251
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORANISM Homo sapiens

```

## REFERENCE

```

1 (bases 1 to 202227)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 202227)
TITLE Direct Submission
AUTHORS Waterston,R.H.
JOURNAL Submitted (28-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On May 4, 2000 this sequence version replaced gi:7658489.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml

```

```

----- Project Information -----
Center project name: H_NH0546114
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-Primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 188247 bases at least Q40
Consensus quality: 193018 bases at least Q30
Consensus quality: 195189 bases at least Q20
Insert size: 174000; agarose-fp
Insert size: 200127; sum-of-contigs
Quality coverage: 4.45 in Q20 bases; agarose-fp
Quality coverage: 3.93 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1412: contig of 1412 bp in length
1413
1512: gap of unknown length
1513
3635: contig of 2123 bp in length
3735: gap of unknown length
3736
4999: contig of 1264 bp in length
5000
5099: gap of unknown length
5100
7667: contig of 2568 bp in length
7668
7767: gap of unknown length
7768
10116: contig of 2349 bp in length
10117
10216: gap of unknown length
10217
14002: contig of 3786 bp in length
14102
14103
14104: gap of unknown length
14105
17164: contig of 3062 bp in length
17165
17264: gap of unknown length
17265
20585: contig of 3321 bp in length
20586
20685: gap of unknown length
20686
24370: contig of 3685 bp in length
24371
24470: gap of unknown length
24471
29951: contig of 5481 bp in length
29952
30051: gap of unknown length
30052
37837: contig of 7786 bp in length
37838
37937: gap of unknown length
37938
48151: contig of 10214 bp in length
48152
48251: gap of unknown length
48252
61495: contig of 13244 bp in length
61496
61595: gap of unknown length
61596
73344: contig of 11749 bp in length
73345
73444: gap of unknown length
73445
85338: contig of 11894 bp in length
85339
85438: gap of unknown length
85439
97336: contig of 11898 bp in length
97337
97436: gap of unknown length
97437
109646: contig of 12210 bp in length
109647
109746: gap of unknown length
109747
123544: contig of 13798 bp in length
123545
123645: gap of unknown length
123646
139857: contig of 16213 bp in length
139858
139957: gap of unknown length
139958
14527: contig of 14570 bp in length
14528
154627: gap of unknown length
154628
174231: contig of 19604 bp in length
174232
174331: gap of unknown length
174332
202227: contig of 27836 bp in length.
Location/Qualifiers
1. 202227
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"

```

```

misc_feature /clone="RP11-546114"
1. 1412
/note="assembly_name:Contig16"
misc_feature 1513. 3635
/note="assembly_name:Contig17"
misc_feature 3736. 4999
/note="assembly_name:Contig18"
misc_feature 5100. 7667
/note="assembly_name:Contig19"
misc_feature 7668. 10116
/note="assembly_name:Contig10"
misc_feature 10217. 14002
/note="assembly_name:Contig11"
14103. 17164
/note="assembly_name:Contig12
clone_end:T7
vector_side:left"
17265. 20585
/note="assembly_name:Contig13"
20686. 24370
/note="assembly_name:Contig14"
24471. 29951
/note="assembly_name:Contig15"
30052. 37837
/note="assembly_name:Contig16"
37938. 48151
/note="assembly_name:Contig17"
48252. 61495
/note="assembly_name:Contig18"
61596. 73344
/note="assembly_name:Contig19"
73445. 85338
/note="assembly_name:Contig20
clone_end:SP6
vector_side:right"
85439. 97336
/note="assembly_name:Contig21"
97437. 109646
/note="assembly_name:Contig22"
109747. 123544
/note="assembly_name:Contig23"
123645. 139857
/note="assembly_name:Contig24"
139958. 154527
/note="assembly_name:Contig25"
154628. 174231
/note="assembly_name:Contig26"
174332. 202227
/note="assembly_name:Contig27"
-----
ORIGIN
Query Match 92.0%; Score 18.4; DB 2; Length 202227;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGACATGCCCGGCGCATGTC 20
Db 184968 GGACATGCCCGGCGCATGTC 184949
-----
RESULT 39
AC087483 203950 bp DNA linear HTG 24-MAY-2001
LOCUS Homo sapiens chromosome 15 clone RP11-546114 map 15, WORKING DRAFT
DEFINITION SEQUENCE, 6 unordered pieces.
AC087483
AC087483.3 GI:14192959
VERSION AC087483.3
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 203950)

```

# AUTHORS TITLE JOURNAL REFERENCE AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 15, clone RP11-546114  
Unpublished  
2 (bases 1 to 203950)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,  
Barna,N., Baetien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,  
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,  
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Karatas,A., Lacroque,K., Lamazares,R., Landers,T.,  
Lecoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,  
McCartney,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,  
McNeters,R., Meldrim,J., Menes,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunhkhang,P., Piere,N., Pollara,V., Raymond,C., Retta,R.,  
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Sougnaz,C., Spencer,B., Strange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,  
Travers,M., Travis,N., Trigliio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (05-JUN-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 24, 2001 this sequence version replaced g1113470198.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: W1BR  
Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L11941

Center clone name: 546\_1\_14

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 202418 bases at least Q40  
Consensus quality: 203020 bases at least Q30  
Consensus quality: 203224 bases at least Q20

Insert size: 205000; agarose-ff  
Insert size: 203450; sum-of-coverage

Quality coverage: 9.5 in Q20 bases; agarose-ff  
Quality coverage: 9.6 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently  
consists of 6 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1701: contig of 1701 bp in length  
\* 1702 1801: gap of 100 bp  
\* 1802 17303: contig of 1508 bp in length  
\* 17310 17409: gap of 100 bp  
\* 17410 60846: contig of 43437 bp in length  
\* 60847 60946: gap of 100 bp  
\* 60947 108903: contig of 47957 bp in length  
\* 108904 109003: gap of 100 bp  
\* 109004 158992: contig of 49889 bp in length  
\* 158993 159092: gap of 100 bp  
\* 159093 203950: contig of 44858 bp in length.

## FEATURES SOURCE

Location/Qualifiers  
1. .203950

## ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 203950;  
Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Ct 1 GGACATGCCCGGCGCATGCC 20  
Db 163747 GGACATGCCCGGCGCATGCC 163766

## RESULT 40 LOCUS DEFINITION

AC087483 203950 bp DNA linear HTG 24-MAY-2001  
Homo sapiens chromosome 15 clone RP11-546114 map 15, WORKING DRAFT  
SEQUENCE, 6 unordered pieces.

AC087483 AC087483.3 GI:14192959  
HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

## SOURCE

ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 203950)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 15, clone RP11-546114  
Unpublished  
2 (bases 1 to 203950)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,  
Barna,N., Baetien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,  
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,  
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Karatas,A., Lacroque,K., Lamazares,R., Landers,T.,  
Lecoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,  
McCartney,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,  
McNeters,R., Meldrim,J., Menes,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunhkhang,P., Piere,N., Pollara,V., Raymond,C., Retta,R.,  
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Sougnaz,C., Spencer,B., Strange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,  
Travers,M., Travis,N., Trigliio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

## JOURNAL REFERENCE AUTHORS

## TITLE

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone="RP11-546114"  
/clone\_lib="RPCI-11 Human Male BAC"  
1. .1701  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
1802. .17309  
/note="assembly\_fragment"  
17410. .60846  
/note="assembly\_fragment"  
60947. .108903  
/note="assembly\_fragment"  
109004. .158992  
/note="assembly\_fragment"  
159093. .203950  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:right"

JOURNAL  
Submitted (05-JUN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 24, 2001 this sequence version replaced gi:13470198.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L11941  
Center clone name: 546.1.14

----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 202418 bases at least Q40  
Consensus quality: 203020 bases at least Q30  
Consensus quality: 203224 bases at least Q20  
Insert size: 205000; agarose-fp  
Insert size: 203450; sum-of-ctnigs  
Quality coverage: 9.5 in Q20 bases; agarose-fp  
Quality coverage: 9.6 in Q20 bases; sum-of-ctnigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      1701: contig of 1701 bp in length
*      1702      1801: gap of 100 bp
*      1802      17309: contig of 15508 bp in length
*      17310      17409: gap of 100 bp
*      17410      60846: contig of 43437 bp in length
*      60847      60946: gap of 100 bp
*      60947      108903: contig of 47957 bp in length
*      108904      109003: gap of 100 bp
*      109004      158992: contig of 49989 bp in length
*      158993      159092: gap of 100 bp
*      159093      203950: contig of 44858 bp in length.

```

FEATURES  
source  
1. .203950  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone\_lib="RP11-546114"  
1. .1701  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"

```

misc_feature      1802..17309  
/note="assembly_fragment"  
misc_feature      17410..60846  
/note="assembly_fragment"  
misc_feature      60947..108903  
/note="assembly_fragment"  
misc_feature      109004..158992  
/note="assembly_fragment"  
misc_feature      159093..203950  
/note="assembly_fragment  
clone_end:T7  
vector_side:right"

```

ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 203950;

Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Ct 1 GGACATGCCCGGCGCATGTCC 20  
Db 163766 GGACATGCTCTGGCATGTCTC 163747

RESULT 41  
AC132452 218688 bp DNA linear ROD 25-NOV-2003  
LOCUS Mus musculus BAC clone RP23-195K10 from chromosome 19, complete  
DEFINITION sequence.  
AC132452  
AC132452.3 GI:33005055  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
TOMLINSON, C. and MEYER, R.  
The sequence of Mus musculus BAC clone RP23-195K10  
Unpublished (2001)  
2 (bases 1 to 218688)  
WILSON, R.  
Sequencing of Mus musculus  
Unpublished (2001)  
3 (bases 1 to 218688)  
McPHERSON, J.D. and WATERSTON, R.H.  
Direct Submission  
Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 218688)  
WILSON, R.K.  
Direct Submission  
Submitted (05-JUN-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 218688)  
WILSON, R.K.  
Direct Submission  
Submitted (20-JUL-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
6 (bases 1 to 218688)  
WILSON, R.  
Direct Submission  
Submitted (25-NOV-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Jul 20, 2003 this sequence version replaced gi:13470198.

REFERENCE  
TOMLINSON, C. and MEYER, R.  
The sequence of Mus musculus BAC clone RP23-195K10  
Unpublished (2001)  
2 (bases 1 to 218688)  
WILSON, R.  
Sequencing of Mus musculus  
Unpublished (2001)  
3 (bases 1 to 218688)  
McPHERSON, J.D. and WATERSTON, R.H.  
Direct Submission  
Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 218688)  
WILSON, R.K.  
Direct Submission  
Submitted (05-JUN-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 218688)  
WILSON, R.K.  
Direct Submission  
Submitted (20-JUL-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
6 (bases 1 to 218688)  
WILSON, R.  
Direct Submission  
Submitted (25-NOV-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Jul 20, 2003 this sequence version replaced gi:13470198.

Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: submissions@wustl.edu  
----- Summary Statistics  
Center project name: M\_BA0195K10

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., paired quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren,

Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

**SOURCE INFORMATION:**

The RPCR-2 BAC library has been constructed by Kazuhiro Oseigawa and Minko Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57Bl/6 mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) Pieter de Jong and coworkers at <http://www.chori.org>

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC133523.

## FEATURES

```

RSES                                     Location/Qualifiers
source                                  1. .218668
                                         /organism="Mus musculus"
                                         /mol_type="genomic DNA"
                                         /db_xref="taxon:10090"
                                         /chromosome="19"
                                         /map="19"
                                         /clone="RP23-195K10"
                                         /clone_11b="RPCT-23"
repeat_region                          291. .427
                                         /rpt_family="Alu"
repeat_region                          444. .629
                                         /rpt_family="B2"
repeat_region                          769. .808
                                         /rpt_family="B2"
repeat_region                          6444. .6561
                                         /rpt_family="B2"
repeat_region                          7006. .7149
                                         /rpt_family="B4"
repeat_region                          7423. .7554
                                         /rpt_family="Alu"
repeat_region                          10624. .10683
                                         /rpt_family="Alu"
repeat_region                          12467. .12653
                                         /rpt_family="B4"
repeat_region                          12472. .12544
                                         /rpt_family="B2"
tRNA                                   /product="tRNA-ser"
                                         /note="Likely pseudogene (HMM Sc=37.16 / Sec struct
                                         Sc=-11.21)"
repeat_region                          13455. .13644
                                         /rpt_family="B2"
repeat_region                          13789. .13923
                                         /rpt_family="Alu"
repeat_region                          13942. .14096
                                         /rpt_family="B2"
repeat_region                          14342. .14460
                                         /rpt_family="B2"
repeat_region                          14594. .14703
                                         /rpt_family="B4"
repeat_region                          16272. .16393
                                         /rpt_family="Alu"
repeat_region                          16796. .16882
                                         /rpt_family="MER1_type"
repeat_region                          16883. .17009
                                         /rpt_family="MER1"
repeat_region                          17738. .17807
                                         /rpt_family="Alu"
repeat_region                          17858. .17964
                                         /rpt_family="ID"
repeat_region                          18293. .18501
                                         /rpt_family="Alu"
repeat_region                          18472. .18537
                                         /rpt_family="B2"
repeat_region                          18507. .18573
                                         /rpt_family="B2"
repeat_region                          18543. .18609
                                         /rpt_family="B2"
repeat_region                          18579. .18645
                                         /rpt_family="B2"

```

```

repeat_region /rpt_family="B2" 18615..18695
repeat_region /rpt_family="B2" 18922..19153
repeat_region /rpt_family="B4" 19341..19482
repeat_region /rpt_family="Alu" 19711..19845
repeat_region /rpt_family="Alu" 20067..20185
repeat_region /rpt_family="Alu" 20277..20423
repeat_region /rpt_family="Alu" 20340..20582
repeat_region /rpt_family="B4" 20708..20824
repeat_region /rpt_family="B4" 21019..21085
repeat_region /rpt_family="Alu" 21225..21271
repeat_region /rpt_family="ERV1" 21272..21410
repeat_region /rpt_family="B4" 21720..21862
repeat_region /rpt_family="HAIIR" 21763..21892
repeat_region /rpt_family="Alu" 22073..22168
repeat_region /rpt_family="Alu" 22242..22348
repeat_region /rpt_family="B4" 22379..22497
repeat_region /rpt_family="Alu" 22819..22962
repeat_region /rpt_family="Alu" 23189..23234
repeat_region /rpt_family="Alu" 23290..23463
repeat_region /rpt_family="B2" 24095..24243
repeat_region /rpt_family="Alu" 24305..24414
repeat_region /rpt_family="Alu" 24449..24631
repeat_region /rpt_family="B2" 24454..24524
tRNA
/product="tRNA-Glu"
/nc="0.04"
repeat_region /rpt_family="Alu" 24838..24947
repeat_region /rpt_family="Alu" 25090..25148
repeat_region /rpt_family="ID" 26564..26610
repeat_region /rpt_family="Alu" 28077..28198
repeat_region /rpt_family="B2" 28235..28372
repeat_region /rpt_family="Alu" 28391..28447
repeat_region /rpt_family="B2" 28476..28657
repeat_region /rpt_family="B2" 28481..28552
tRNA

```

```
QY      1 GGACATGCCCGGCATGTCC 20
         |||||
Db       60291 GGACTGCCCCGGCATGTCC 60310
          | | | | | | | | | |
Query Match Score 18.4; DB 10; Length 216688;
Beet Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

RESULT 42  
AC132452/c 218688 bp DNA linear ROD 25-NOV-2003  
LOCUS AC132452  
DEFINITION Mus musculus BAC clone RP23-195K10 from chromosome 19, complete  
sequence.  
AC132452.3 GI:33005055  
HTG.  
MUS musculus (house mouse)  
MUS musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 218688)  
Tomlinson, C. and Meyer, R.  
The sequence of Mus musculus BAC clone RP23-195K10  
Unpublished (2001)  
2 (bases 1 to 218688)  
Wilson, R.  
Sequencing of Mus musculus  
Unpublished (2001)  
3 (bases 1 to 218688)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 218688)  
Wilson, R.K.  
Direct Submission  
Submitted (05-JUN-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 218688)  
Wilson, R.K.  
Direct Submission  
Submitted (20-JUL-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
6 (bases 1 to 218688)  
Wilson, R.  
Direct Submission  
Submitted (25-NOV-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Jul 20, 2003 this sequence version replaced gi:31416109.  
-----  
Center: Washington University Genome Sequencing Center  
Genome Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
-----  
Summary Statistics  
Center project name: M\_BA0195K10  
-----

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
<http://genome.wustl.edu>

SOURCE INFORMATION:  
The RPCT-23 BAC Library has been constructed by Kazutoyo Oseegawa  
and Minako Tateno in the laboratory of Pieter de Jong

# FEATURES

source  
  
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or  
brain genomic DNA. The clone and detailed information can be  
obtained from Research Genetics, Inc. (<http://www.reagen.com>) or  
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:  
This sequence is the entire insert of the clone. This clone is  
overlapped by AC133523.

## Location/Qualifiers

```
1..218688
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="19"
/map="19"
/clone="RP23-195K10"
/clone_1bp="RPCT-23"
291..427
/rpt_family="Alu"
444..629
/rpt_family="B2"
769..808
/rpt_family="B2"
6444..6561
/rpt_family="B4"
7006..7149
/rpt_family="Alu"
7423..7554
/rpt_family="Alu"
10624..10683
/rpt_family="B4"
12467..12653
/rpt_family="B2"
12472..12544
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=37.16 / Sec struct
Sc=-11.21)"
13455..13644
/rpt_family="B2"
13789..13923
/rpt_family="Alu"
13942..14096
/rpt_family="B2"
14342..14460
/rpt_family="B4"
14594..14703
/rpt_family="Alu"
16272..16393
/rpt_family="MER1_type"
16796..16882
/rpt_family="MER1_type"
16883..17009
/rpt_family="Alu"
17738..17807
/rpt_family="ID"
17858..17964
/rpt_family="Alu"
18293..18501
/rpt_family="B2"
18472..18537
/rpt_family="B2"
18507..18573
/rpt_family="B2"
18543..18609
/rpt_family="B2"
18579..18645
/rpt_family="B2"
18615..18695
/rpt_family="B2"
18922..19153
/rpt_family="B4"
19341..19482
/rpt_family="Alu"
```



```

repeat_region      19711..19845
                    /rpt_family="Alu"
repeat_region      20667..20185
                    /rpt_family="Alu"
repeat_region      20277..20423
                    /rpt_family="Alu"
repeat_region      20540..20582
                    /rpt_family="B4"
repeat_region      20708..20824
                    /rpt_family="B4"
repeat_region      21019..21085
                    /rpt_family="Alu"
repeat_region      21225..21271
                    /rpt_family="ERV1"
repeat_region      21272..21410
                    /rpt_family="B4"
repeat_region      21720..21862
                    /rpt_family="MALR"
repeat_region      21762..21892
                    /rpt_family="Alu"
repeat_region      22073..22168
                    /rpt_family="Alu"
repeat_region      22242..22348
                    /rpt_family="B4"
repeat_region      22379..22497
                    /rpt_family="Alu"
repeat_region      22819..22962
                    /rpt_family="Alu"
repeat_region      23189..23234
                    /rpt_family="Alu"
repeat_region      23290..23463
                    /rpt_family="B2"
repeat_region      24095..24243
                    /rpt_family="Alu"
repeat_region      24305..24414
                    /rpt_family="Alu"
repeat_region      24449..24631
                    /rpt_family="B2"
repeat_region      24451..24524
                    /product="tRNA-Glu"
                    /note="Likely pseudogene (HMM Sc=23.47 / Sec struct
                    Sc=0.04)"
repeat_region      24838..24947
                    /rpt_family="Alu"
repeat_region      25090..25148
                    /rpt_family="ID"
repeat_region      26564..26610
                    /rpt_family="Alu"
repeat_region      28077..28198
                    /rpt_family="B2"
repeat_region      28235..28372
                    /rpt_family="Alu"
repeat_region      28391..28447
                    /rpt_family="B2"
repeat_region      28476..28657
                    /rpt_family="B2"
repeat_region      28481..28552
                    /rpt_family="B2"
tRNA

```

Query Match 92.0%; Score 18.4; DB 10; Length 218688;  
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GCACATGCCGGCATGTCC 20  
 60310 GCACATGCCGGCATGTCC 60291

RESULT 43  
 AC113311 260120 bp DNA linear ROD 29-JAN-2004  
 LOCUS AC113311 Mus musculus chromosome 1, clone RP23-440J24, complete sequence.  
 DEFINITION AC113311  
 ACCESSION AC113311.9 GI:41386878  
 VERSION

KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

HTG.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 260120)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Mus musculus chromosome 1, clone RP23-440J24  
 Unpublished  
 2 (bases 1 to 260120)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fazo,S.,  
 Ferrreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Glinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Labrecque,K., Lamazares,R.,  
 Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,  
 Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M.,  
 McEwan,P., McKernan,K., Meldrim,J., Menus,L., Mihova,T.,  
 Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,  
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,  
 Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,  
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S.,  
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travi,N., Triglio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (28-FEB-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 260120)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Corum,B., Dekrellano,K.,  
 Diaz,J.S., Dodge,S., Doolley,K., Dorris,L., Erickson,J., Fazo,S.,  
 Ferrreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hago,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,  
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrim,J., Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P.,  
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbbs,M.,  
 Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,  
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 260120)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Corum,B., Dekrellano,K.,  
 Diaz,J.S., Dodge,S., Doolley,K., Dorris,L., Erickson,J., Fazo,S.,  
 Ferrreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hago,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,  
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrim,J., Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,

TITLE  
JOURNAL

## COMMENT

O'Neil, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vasilev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (29-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 29, 2004 this sequence version replaced gi:3979533.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBK  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
Project Information  
Center project name: L22773  
Center clone name: 440\_J\_24

## FEATURES

## source

## misc\_feature

Location/Qualifiers  
1..260120  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"  
/clone="RP23-440J24"  
/clone\_id="RPCI-23 Female Mouse BAC"  
1..37846  
/note="wgs end extension  
clone end:5P6"  
repeat\_region  
131..178  
/rpt\_family="(TG)n"  
repeat\_region  
180..313  
/rpt\_family="(L2)n"  
repeat\_region  
974..1253  
/rpt\_family="(RMR6B)  
complement(11327..1727)  
repeat\_region  
1733..1934  
/rpt\_family="(Lx5"  
repeat\_region  
2563..2598  
/rpt\_family="(RMR6B"  
repeat\_region  
complement(12589..2742)  
/rpt\_family="(B1\_MM"  
repeat\_region  
complement(2940..3147)  
/rpt\_family="(L1MA4A"  
repeat\_region  
3780..3810  
/rpt\_family="(T)n"  
complement(3891..4578)  
repeat\_region  
4606..4690  
/rpt\_family="(B4A"  
repeat\_region  
4935..4958  
/rpt\_family="(AT\_rich"  
repeat\_region  
5026..5089  
/rpt\_family="(AT\_rich"  
repeat\_region  
5702..6084  
/rpt\_family="(MTD"  
complement(6304..6426)  
repeat\_region  
6488..6589  
/rpt\_family="(MIR"  
complement(6964..7174)  
repeat\_region  
7206..7263  
/rpt\_family="(7SK"  
repeat\_region  
7849..7873  
/rpt\_family="(Charlela"  
/rpt\_family="(TG)n"

repeat\_region 7873..7908  
/rpt\_family="(GA)n"  
complement(9173..9698)  
repeat\_region  
/rpt\_family="(Lx"  
complement(9791..9848)  
repeat\_region  
/rpt\_family="(B3"  
9971..10126  
/rpt\_family="(MIR"  
complement(10237..10433)  
repeat\_region  
/rpt\_family="(B3A"  
10849..11080  
/rpt\_family="(B4"  
complement(11170..11371)  
repeat\_region  
/rpt\_family="(CA)n"  
complement(11515..11626)  
repeat\_region  
/rpt\_family="(Looper"  
11713..11743  
complement(11922..12107)  
repeat\_region  
/rpt\_family="(Looper"  
12108..12238  
/rpt\_family="(RSINE1"  
12239..12275  
/rpt\_family="(TG)n"  
repeat\_region  
12686..12738  
/rpt\_family="(AT\_rich"  
13220..13240  
/rpt\_family="(AT\_rich"  
13511..13538  
/rpt\_family="(AT\_rich"  
13548..13696  
/rpt\_family="(RSINE1"  
14736..14757  
/rpt\_family="(AT\_rich"  
complement(14930..14998)  
/note="single clone coverage"  
complement(15003..15085)  
repeat\_region  
/rpt\_family="(PB1D10"  
complement(15450..15640)  
repeat\_region  
/rpt\_family="(B2\_Mm2"  
complement(15670..15904)  
repeat\_region  
/rpt\_family="(L1"  
16334..16373  
/rpt\_family="(AT\_rich"  
complement(16493..16674)  
repeat\_region  
/rpt\_family="(ORR1D"  
16990..17154  
/rpt\_family="(URR1A"  
18058..18109  
repeat\_region  
/rpt\_family="(AT\_rich"

Query Match 92.0%; Score 18.4; DB 10; Length 260120;

Best Local Similarity 95.0%; Pred.No.2.2e+02; Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
DB 81584 GGACATGCCCGGCGCATGTCC 81603

RESULT 44  
AC113311 260120 bp DNA linear ROD 29-JAN-2004  
LOCUS Mus musculus chromosome 1, clone RP23-440J24, complete sequence.  
DEFINITION  
AC113311 AC113311.9 GI:41386678  
VERSION  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 260120)  
Birren, B., Nusbaum, C. and Lander, B.



```

repeat_region      /rpt_family="MIR"
                    complement(10237..10433)
repeat_region      /rpt_family="B3A"
                    10849..11080
repeat_region      /rpt_family="B4"
                    complement(11170..11371)
repeat_region      /rpt_family="looper"
                    complement(11515..11626)
repeat_region      /rpt_family="looper"
                    11713..11743
repeat_region      /rpt_family="(CA)n"
                    complement(11922..12107)
repeat_region      /rpt_family="looper"
                    12108..12238
repeat_region      /rpt_family="RSINE1"
                    12239..12275
repeat_region      /rpt_family="(TG)n"
                    12686..12738
repeat_region      /rpt_family="AT_rich"
                    13220..13240
repeat_region      /rpt_family="AT_rich"
                    13511..13538
repeat_region      /rpt_family="AT_rich"
                    13548..13696
repeat_region      /rpt_family="RSINE1"
                    14736..14757
repeat_region      /rpt_family="AT_rich"
                    complement(14930..14998)
unsure             /note="single clone coverage"
repeat_region      complement(15003..15085)
repeat_region      /rpt_family="PB1D10"
                    complement(15450..15640)
repeat_region      /rpt_family="B2_Mm2"
                    complement(15670..15904)
repeat_region      /rpt_family="L1"
                    16334..16373
repeat_region      /rpt_family="AT_rich"
                    complement(16493..16674)
repeat_region      /rpt_family="ORR1D"
                    16980..17154
repeat_region      /rpt_family="ORR1A"
                    18058..18109
repeat_region      /rpt_family="AT_rich"

```

Query Match 92.0%; Score 18.4; DB 10; Length 260120;  
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGCGCATGTCC 20  
 DB 81603 GGACCTGCCCGCATGTCC 81584

RESULT 45  
 AL807771 233077 bp DNA linear ROD 26-NOV-2002  
 LOCUS Mouse DNA sequence from clone RP23-141C15 on chromosome 4, complete  
 DEFINITION  
 ACCESSION AL807771  
 VERSION AL807771.6 GI:25809489  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Lovell, J.  
 TITLE Direct Submission  
 AUTHORS Submitted (26-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,  
 JOURNAL Cambridgehire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Nov 27, 2002 this sequence version replaced gi:21912779.  
 ----- Genome Center

Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; information on the WORMPEP database can be found at [http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) RP23-141C15 is from the RP23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.choxi.org/bacpac/home.htm> VECTOR: PBAC3.6

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

## FEATURES

location/Qualifiers  
 1..233077  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-141C15"  
 /clone\_1db="RP23-23"

## ORIGIN

Query Match 90.0%; Score 18; DB 10; Length 233077;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGCGCATGT 18  
 DB 193609 GGACATGCCCGCGCATGT 193626

RESULT 46  
 AL807771/c 233077 bp DNA linear ROD 26-NOV-2002  
 LOCUS Mouse DNA sequence from clone RP23-141C15 on chromosome 4, complete  
 DEFINITION  
 ACCESSION AL807771  
 VERSION AL807771.6 GI:25809489  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Lovell, J.  
 TITLE Direct Submission  
 AUTHORS Submitted (26-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,  
 JOURNAL Cambridgehire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Nov 27, 2002 this sequence version replaced gi:21912779.  
 ----- Genome Center  
 Center code: SC

Web site: <http://www.sanger.ac.uk>  
Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-141C15 is from the RPI-23 Mouse PAC library.

constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

## FEATURES

## source

1. 233077  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone="RP23-141C15"  
/clone\_1fb="RPI-23"  
/clone\_1lb="RPI-23"

## ORIGIN

Query Match 90.0%; Score 18; DB 10; Length 233077;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ACATGCCCCGGCATGTCC 20  
DB 193626 ACATGCCCCGGCATGTCC 193609

## RESULT 47

## AC093466

LOCUS AC093466 263507 bp DNA linear HTG 23-APR-2003  
DEFINITION Mus musculus clone RP23-4M7, WORKING DRAFT SEQUENCE, 11 unordered pieces.

## AC093466

VERSION AC093466.2 GI:30018208  
HTG, HTGS PHASE1, HTGS\_DRAFT.

KEYWORDS Mus musculus (house mouse)

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Unpublished  
(bases 1 to 263507)  
Birren, B., Nussbaum, C. and Lander, E.  
2 (bases 1 to 263507)  
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgibbon, W., Gage, D., Galagan, J., Gardyna, S., Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamarche, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McMan, P., McKernan, K., McPheters, R., Melim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollard, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strassman, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (26-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 263507)  
Birren, B., Nussbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Cornu, B., Dearellano, K., Diaz, J. S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgibbon, W., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 17, 2003 this sequence version replaced gi:15290913.  
All repeats were identified using RepeatMasker:  
Smit, A. P. A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L13364  
Center clone name: 4\_M\_7

\*\*\*\*\* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 50429: contig of 50429 bp in length  
\* 50430 50529: gap of 100 bp  
\* 50530 52201: contig of 1672 bp in length  
\* 52202 52301: gap of 100 bp  
\* 52302 54425: contig of 2124 bp in length  
\* 54426 54525: gap of 100 bp  
\* 54526 57746: contig of 3221 bp in length  
\* 57747 57847: gap of 100 bp  
\* 57847 60652: contig of 2806 bp in length  
\* 60653 60752: gap of 100 bp

```

* 60753 64812: contig of 4060 bp in length
* 64813 64912: gap of 100 bp
* 64913 79944: contig of 15032 bp in length
* 79945 80044: gap of 100 bp
* 80045 101704: contig of 21660 bp in length
* 101705 101804: gap of 100 bp
* 101805 169526: contig of 67722 bp in length
* 169527 203765: gap of 100 bp
* 203766 203856: contig of 34139 bp in length
* 203857 263507: contig of 100 bp
* 203858 263507: contig of 59642 bp in length.

```

## FEATURES

## source

```

1. 263507
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-4M7"
/clone_lib="RP23-4M7"
/clone_id="RP23-4M7"
1. 50429
/clone="assembly_fragment"

```

## misc\_feature

```

clone_end:SP6
vector_side:left"

```

## misc\_feature

```

50530..52201
/clone="assembly_fragment"

```

## misc\_feature

```

52302..54425
/clone="assembly_fragment"

```

## misc\_feature

```

54526..57746
/clone="assembly_fragment"

```

## misc\_feature

```

57847..60652
/clone="assembly_fragment"

```

## misc\_feature

```

60753..64812
/clone="assembly_fragment"

```

## misc\_feature

```

64913..79944
/clone="assembly_fragment"

```

## misc\_feature

```

80045..101704
/clone="assembly_fragment"

```

## misc\_feature

```

101805..169526
/clone="assembly_fragment"

```

## misc\_feature

```

169627..203765
/clone="assembly_fragment"

```

## misc\_feature

```

203856..263507
/clone="assembly_fragment"

```

```

clone_end:77
vector_side:right"

```

## ORIGIN

## Query Match

```

Best local similarity 90.0%; Score 18; DB 2; Length 263507;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

3 ACATGCCCGCGCATGCC 20
|||||

```

```

DB 197519 ACATGCCCGCGCATGCC 197536

```

## RESULT 48

## AC093466/c

## LOCUS

## DEFINITION

## AC093466

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

```

AC093466 263507 bp DNA linear HTG 23-APR-2003
Mus musculus clone RP23-4M7, WORKING DRAFT SEQUENCE, 11 unordered
pieces.
AC093466 GI:30018208
HTG; HTGS PHASE1; HTGS DRAFT.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 263507)
Birren,B., Nuebaum,C. and Lander,E.
Unpublished
2 (bases 1 to 263507)
Birren,B., Linton,L., Nuebaum,C., Lander,E., Ali,A., Allen,N.,

```

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Anderson, S., Barna, N., Baetien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeBellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gattada, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Marguis, N., Matthews, C., McClean, M., McEwan, P., McKernan, K., McNetters, R., Meldrum, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, N., Schauback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (26-AUG-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 263507)

Birren, B., Nuebaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choquet, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeBellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, N., Gage, D., Galagan, J., Gattada, S., Graham, L., Grand-Pierre, N., Hages, B., Hales, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., McClean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, N., Schauback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

## Genome Center

## Center: Whitehead Institute/ MIT Center for Genome Research

## Center code: WIBR

## Web site: http://www-seq.wi.mit.edu

## Contact: sequence\_submissions@genome.wi.mit.edu

## Project Information

## Center project name: L13564

## Center clone name: 4\_M\_7

NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 50429: contig of 50429 bp in length  
\* 50430 50529: gap of 100 bp  
\* 50530 52201: contig of 1672 bp in length  
\* 52202 52301: gap of 100 bp

```

* 52302 54425: contig of 2124 bp in length
* 54426 54525: gap of 100 bp
* 54526 57746: contig of 3221 bp in length
* 57747 57846: gap of 100 bp
* 57847 60652: contig of 2806 bp in length
* 60653 60752: gap of 100 bp
* 60753 64812: contig of 4060 bp in length
* 64813 64912: gap of 100 bp
* 64913 79944: contig of 15032 bp in length
* 79945 80044: gap of 100 bp
* 80045 101704: contig of 21660 bp in length
* 101705 101804: gap of 100 bp
* 101805 169526: contig of 67722 bp in length
* 169527 169627: gap of 100 bp
* 169627 203765: contig of 34139 bp in length
* 203766 203865: gap of 100 bp
* 203866 263507: contig of 59642 bp in length.
FEATURES
source
1..263507
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_id="RP23-4M7"
/clone_lib="RPC1-23 Female Mouse BAC"
1..50429
/clone_end:SP6
/vector_side:left"
misc_feature
50530..52201
/note="assembly_fragment"
52302..54425
/note="assembly_fragment"
54526..57746
/note="assembly_fragment"
57847..60652
/note="assembly_fragment"
60753..64812
/note="assembly_fragment"
64913..79944
/note="assembly_fragment"
80045..101704
/note="assembly_fragment"
101805..169526
/note="assembly_fragment"
169627..203765
/note="assembly_fragment"
203866..263507
/note="assembly_fragment"
clone_end:T7
vector_side:right"

```

## ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 263507;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGT 18  
 |||||  
 Db 197536 GGACATGCCCGGCATGT 197519

## RESULT 49

AY395415 1439 bp DNA linear BCT 16-JUN-2004  
 LOCUS AY395415  
 DEFINITION Uncultured Rubrobacteridae bacterium clone EBI096 16S ribosomal RNA  
 ACCESSION AY395415  
 VERSION AY395415.1 GI:39546077  
 KEYWORDS  
 SOURCE uncultured Rubrobacteridae bacterium  
 ORGANISM Bacteria; Actinobacteria; Rubrobacteridae; environmental samples.  
 REFERENCE 1 (bases 1 to 1439)

AUTHORS Schoenborn, L., Salt, M., Hugenholz, P. and Janssen, P.H.  
 TITLE Major groups of bacteria in a pasture soil  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1439)  
 AUTHORS Schoenborn, L., Salt, M., Hugenholz, P. and Janssen, P.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-SEP-2003) Department of Microbiology and Immunology,  
 University of Melbourne, Grattan Street, Parkville, Victoria 3010,  
 Australia  
FEATURES  
source  
1..1439  
/organism="uncultured Rubrobacteridae bacterium"  
/mol\_type="genomic DNA"  
/isolation\_source="pasture soil"  
/db\_xref="taxon:211429"  
/clone="EBI096"  
/environmental\_sample  
<1..>1439  
/product="16S ribosomal RNA"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 1; Length 1439;  
 Best Local Similarity 94.7%; Pred. No. 9.6e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGACATGCCCGGCATGTC 20  
 |||||  
 Db 969 GGACATGCCCGGCATGTC 987

RESULT 50  
 AY395415 1439 bp DNA linear BCT 16-JUN-2004  
 LOCUS AY395415  
 DEFINITION Uncultured Rubrobacteridae bacterium clone EBI096 16S ribosomal RNA  
 ACCESSION AY395415  
 VERSION AY395415.1 GI:39546077  
 KEYWORDS  
 SOURCE uncultured Rubrobacteridae bacterium  
 ORGANISM Bacteria; Actinobacteria; Rubrobacteridae; environmental samples.  
 REFERENCE 1 (bases 1 to 1439)  
 AUTHORS Schoenborn, L., Salt, M., Hugenholz, P. and Janssen, P.H.  
 TITLE Major groups of bacteria in a pasture soil  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1439)  
 AUTHORS Schoenborn, L., Salt, M., Hugenholz, P. and Janssen, P.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-SEP-2003) Department of Microbiology and Immunology,  
 University of Melbourne, Grattan Street, Parkville, Victoria 3010,  
 Australia  
FEATURES  
source  
1..1439  
/organism="uncultured Rubrobacteridae bacterium"  
/mol\_type="genomic DNA"  
/isolation\_source="pasture soil"  
/db\_xref="taxon:211429"  
/clone="EBI096"  
/environmental\_sample  
<1..>1439  
/product="16S ribosomal RNA"

## FEATURES

Query Match 87.0%; Score 17.4; DB 1; Length 1439;  
 Best Local Similarity 94.7%; Pred. No. 9.6e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## ORIGIN

Qy 1 GGACATGCCCGGCATGTC 19  
 |||||  
 Db 987 GGACATGCCCGGCATGTC 969

Search completed: January 14, 2005, 16:03:44

Job time : 1050.74 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:09:04 ; Search time 334.737 Seconds  
(without alignments)  
313.645 Million cell updates/sec

Title: US-09-578-453-2  
Perfect score: 20  
Sequence: 1 GGACATGCCCGGCATGTC 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 413486 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database :  
1: N\_Geneseq\_23Sep04:\*  
2: geneeqn1980s:\*  
3: geneeqn1990s:\*  
4: geneeqn2000s:\*  
5: geneeqn2001s:\*  
6: geneeqn2002s:\*  
7: geneeqn2002bs:\*  
8: geneeqn2003s:\*  
9: geneeqn2003bs:\*  
10: geneeqn2003cs:\*  
11: geneeqn2003ds:\*  
12: geneeqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAQ88279	p53 prote
2	20	100.0	20	AAQ88279	p53 prote
3	20	100.0	20	AAV22590	p53 conse
4	20	100.0	20	AAV22590	p53 conse
5	20	100.0	20	AAV22590	p53 conse
6	20	100.0	20	AAV22590	p53 conse
7	20	100.0	20	AAV22590	p53 conse
8	20	100.0	20	AAV22590	p53 conse
9	20	100.0	20	AAV22590	p53 conse
10	20	100.0	20	AAV22590	p53 conse
11	20	100.0	20	AAV22590	p53 conse
12	20	100.0	20	AAV22590	p53 conse
13	20	100.0	20	AAV22590	p53 conse
14	20	100.0	20	AAV22590	p53 conse
15	20	100.0	20	AAV22590	p53 conse
16	20	100.0	20	AAV22590	p53 conse
17	20	100.0	20	AAV22590	p53 conse
18	20	100.0	20	AAV22590	p53 conse
19	20	100.0	20	AAV22590	p53 conse
20	20	100.0	20	AAV22590	p53 conse
21	20	100.0	20	AAV22590	p53 conse

22	20	100.0	27	3	AAA61640	AAA61640 p53 bindi
23	20	100.0	56	3	AAA38055	AAA38055 linker us
24	20	100.0	56	3	AAA38055	AAA38055 linker us
25	20	100.0	56	8	ABX16465	ABX16465 selective
26	20	100.0	56	8	ABX16465	ABX16465 selective
27	20	100.0	56	8	ABX16450	ABX16450 p53CON-1u
28	20	100.0	56	8	ABX16450	ABX16450 p53CON-1u
29	20	100.0	56	8	ABX16449	ABX16449 p53CON-1u
30	20	100.0	56	8	ABX16449	ABX16449 p53CON-1u
31	20	100.0	56	8	ABX16449	ABX16449 p53CON-1u
32	20	100.0	56	8	ABX16466	ABX16466 selective
33	20	100.0	56	8	ABX16466	ABX16466 selective
34	19	95.0	19	12	ADBS2736	ADBS2736 Oligonuc
35	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
36	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
37	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
38	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
39	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
40	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
41	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
42	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
43	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
44	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
45	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
46	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
47	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
48	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
49	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
50	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
51	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
52	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
53	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
54	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
55	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
56	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
57	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
58	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
59	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
60	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
61	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
62	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
63	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
64	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
65	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
66	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
67	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
68	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
69	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
70	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
71	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
72	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
73	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
74	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
75	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
76	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
77	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
78	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
79	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
80	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
81	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
82	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
83	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
84	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
85	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
86	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
87	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
88	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
89	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
90	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
91	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
92	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
93	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su
94	16.8	84.0	30	2	AAAT32849	AAAT32849 Tumour su

95	15.4	77.0	473	6	ABL34893	Abi34893 Murine cd	c 168	15.2	76.0	1920	6	AA29931	AA29931 Human cal
96	15.4	77.0	473	6	ABL34893	Abi34893 Murine cd	c 169	15.2	76.0	1924	6	AA95010	AA95010 Human DNA
97	15.4	77.0	7282	10	ACC4359	Acc43595 Human NR1	c 170	15.2	76.0	1924	6	AA95010	AA95010 Human DNA
98	15.4	77.0	7282	10	ACC4359	Acc43595 Human NR1	c 171	15.2	76.0	1940	9	ADA2636	ADA2636 Human cal
99	15.4	77.0	114411	12	ADQ21090	Adq21090 Human sof	c 172	15.2	76.0	1940	9	ADA2636	ADA2636 Human cal
100	15.4	77.0	114411	12	ADQ21090	Adq21090 Human sof	c 173	15.2	76.0	1958	3	AAA09346	AAA09346 Human MBP
101	15.2	76.0	20	2	AAVI0231	Aavi0231 p53 bindi	c 174	15.2	76.0	1958	3	AAA09346	AAA09346 Human MBP
102	15.2	76.0	20	2	AAVI0231	Aavi0231 p53 bindi	c 175	15.2	76.0	1958	3	AAA09346	AAA09346 Human MBP
103	15.2	76.0	24	12	ADP26898	Adp26898 Transcrip	c 176	15.2	76.0	1958	6	AA239469	AA239469 Human cal
104	15.2	76.0	24	12	ADP26898	Adp26898 Transcrip	c 177	15.2	76.0	1958	6	AA239469	AA239469 Human cal
105	15.2	76.0	291	3	AACT18258	Aact18258 Human sec	c 178	15.2	76.0	1958	6	ADG34537	ADG34537 Human cal
106	15.2	76.0	291	3	AACT18258	Aact18258 Human sec	c 179	15.2	76.0	2934	9	ADA48441	ADA48441 Rice gene
107	15.2	76.0	297	6	ABK27635	Abk27635 Human col	c 180	15.2	76.0	2934	9	ADA48441	ADA48441 Rice gene
108	15.2	76.0	297	6	ABK27635	Abk27635 Human col	c 181	15.2	76.0	4067	3	AAA95823	AAA95823 Human met
109	15.2	76.0	300	6	ABL37865	Abi37865 Human col	c 182	15.2	76.0	4067	3	AAA95823	AAA95823 Human met
110	15.2	76.0	300	6	ABL37865	Abi37865 Human col	c 183	15.2	76.0	4258	8	ABT18136	ABT18136 Aspergill
111	15.2	76.0	366	2	AAK27251	Aak27251 Human col	c 184	15.2	76.0	4258	8	ABT18136	ABT18136 Aspergill
112	15.2	76.0	366	2	AAK27251	Aak27251 Human col	c 185	15.2	76.0	4258	8	ABT19950	ABT19950 Aspergill
113	15.2	76.0	401	9	ACH17140	Ach17140 Human adu	c 186	15.2	76.0	4258	8	ABT19950	ABT19950 Aspergill
114	15.2	76.0	401	9	ACH17140	Ach17140 Human adu	c 187	15.2	76.0	5734	10	ADC85382	ADC85382 Human ptp
115	15.2	76.0	442	2	AAV89143	Aav89143 EST clone	c 188	15.2	76.0	5734	10	ADC85382	ADC85382 Human ptp
116	15.2	76.0	442	2	AAV89143	Aav89143 EST clone	c 189	15.2	76.0	6942	2	AAQ86480	AAQ86480 Merosin p
117	15.2	76.0	463	10	ADF85659	Adf85659 Human ade	c 190	15.2	76.0	6942	2	AAQ86480	AAQ86480 Merosin p
118	15.2	76.0	463	10	ADF85659	Adf85659 Human ade	c 191	15.2	76.0	9419	3	AAA88894	AAA88894 Human lam
119	15.2	76.0	472	9	ACH25933	Ach25933 Human adu	c 192	15.2	76.0	9419	3	AAA88894	AAA88894 Human lam
120	15.2	76.0	472	9	ACH25933	Ach25933 Human adu	c 193	15.2	76.0	9420	3	AAA88892	AAA88892 Human lam
121	15.2	76.0	491	9	ACH23918	Ach23918 Human adu	c 194	15.2	76.0	9420	3	AAA88892	AAA88892 Human lam
122	15.2	76.0	491	9	ACH23918	Ach23918 Human adu	c 195	15.2	76.0	9534	2	AAI17419	AAI17419 Merosin c
123	15.2	76.0	492	9	ACH14709	Ach14709 Human adu	c 196	15.2	76.0	9534	2	AAI17419	AAI17419 Merosin c
124	15.2	76.0	492	9	ACH14709	Ach14709 Human adu	c 197	15.2	76.0	9534	2	AAK59768	AAK59768 CDNA enco
125	15.2	76.0	538	12	ACH80089	Ach80089 Human gen	c 198	15.2	76.0	9534	2	AAK59768	AAK59768 CDNA enco
126	15.2	76.0	538	12	ACH80089	Ach80089 Human gen	c 199	15.2	76.0	9534	3	AAA88893	AAA88893 Human lam
127	15.2	76.0	569	12	ACH80444	Ach80444 Human gen	c 200	15.2	76.0	9534	3	AAA88893	AAA88893 Human lam
128	15.2	76.0	569	12	ACH80444	Ach80444 Human gen	c 201	15.2	76.0	9534	3	AAA88893	AAA88893 Human lam
129	15.2	76.0	574	12	ACH73863	Ach73863 Human gen	c 202	15.2	76.0	9534	6	ABL64110	ABL64110 Breast ca
130	15.2	76.0	574	12	ACH73863	Ach73863 Human gen	c 203	15.2	76.0	9534	6	ABK35565	ABK35565 Gene LAMA
131	15.2	76.0	615	10	ABT3245	Abt3245 Seed deve	c 204	15.2	76.0	9534	6	ABK35565	ABK35565 Gene LAMA
132	15.2	76.0	615	10	ABT3245	Abt3245 Seed deve	c 205	15.2	76.0	9534	6	ABT11110	ABT11110 Human bre
133	15.2	76.0	1024	10	ABZ83272	Abz83272 Toxicolog	c 206	15.2	76.0	9534	6	ABT11110	ABT11110 Human bre
134	15.2	76.0	1024	10	ABZ83272	Abz83272 Toxicolog	c 207	15.2	76.0	9534	10	ADK61038	ADK61038 Ovarian c
135	15.2	76.0	1152	8	ADA70053	Ada70053 Rice gene	c 208	15.2	76.0	9534	10	ADK61038	ADK61038 Ovarian c
136	15.2	76.0	1152	8	ADA70053	Ada70053 Rice gene	c 209	15.2	76.0	9535	3	AAA88891	AAA88891 Human lam
137	15.2	76.0	1152	11	ADU11565	Adj11565 Rice DNA	c 210	15.2	76.0	9535	3	AAA88891	AAA88891 Human lam
138	15.2	76.0	1152	11	ADU11565	Adj11565 Rice DNA	c 211	15.2	76.0	9649	12	ADQ17880	ADQ17880 Human sof
139	15.2	76.0	1251	3	AAA09347	Aaa09347 Human MBP	c 212	15.2	76.0	9649	12	ADQ17880	ADQ17880 Human sof
140	15.2	76.0	1251	3	AAA09347	Aaa09347 Human MBP	c 213	15.2	76.0	9670	12	ADQ22504	ADQ22504 Human sof
141	15.2	76.0	1434	5	AAS91386	Aas91386 DNA enco	c 214	15.2	76.0	9670	12	ADQ22504	ADQ22504 Human sof
142	15.2	76.0	1434	5	AAS91386	Aas91386 DNA enco	c 215	15.2	76.0	9746	5	AA84601	AA84601 DNA enco
143	15.2	76.0	1436	5	AAS67726	Aas67726 DNA enco	c 216	15.2	76.0	9746	5	AA84601	AA84601 DNA enco
144	15.2	76.0	1436	5	AAS67726	Aas67726 DNA enco	c 217	15.2	76.0	10867	4	AA836613	AA836613 Human car
145	15.2	76.0	1481	10	ADG75777	Adg75777 Human pro	c 218	15.2	76.0	10867	4	AA836613	AA836613 Human car
146	15.2	76.0	1481	10	ADG75777	Adg75777 Human pro	c 219	15.2	76.0	10867	4	AA836613	AA836613 Human car
147	15.2	76.0	1890	1	AAN91471	Aan91471 1.9 kb Ro	c 220	15.2	76.0	10867	4	AA836613	AA836613 Human car
148	15.2	76.0	1890	1	AAN91471	Aan91471 1.9 kb Ro	c 221	15.2	76.0	10867	4	AA836613	AA836613 Human car
149	15.2	76.0	1890	10	ADD22456	Add22456 HLA-B46 T	c 222	15.2	76.0	10867	4	AA836613	AA836613 Human car
150	15.2	76.0	1890	10	ADD22456	Add22456 HLA-B46 T	c 223	15.2	76.0	10867	4	AA836613	AA836613 Human car
151	15.2	76.0	1899	6	ABK11662	Abk11662 DNA enco	c 224	15.2	76.0	10867	8	ABX60593	ABX60593 CDNA enco
152	15.2	76.0	1899	6	ABK11662	Abk11662 DNA enco	c 225	15.2	76.0	10867	8	ABX60593	ABX60593 CDNA enco
153	15.2	76.0	1899	6	ABV73173	Abv73173 Human cal	c 226	15.2	76.0	10867	10	ADB47307	ADB47307 Human car
154	15.2	76.0	1899	6	ABV73173	Abv73173 Human cal	c 227	15.2	76.0	10867	12	ADJ31343	ADJ31343 Human mus
155	15.2	76.0	1899	10	ADF81442	Adf81442 Leukaemia	c 228	15.2	76.0	10867	12	ADJ31343	ADJ31343 Human mus
156	15.2	76.0	1899	10	ADF81442	Adf81442 Leukaemia	c 229	15.2	76.0	10894	4	AA836611	AA836611 Human car
157	15.2	76.0	1899	10	ADIS7348	Adi57348 Human pyu	c 230	15.2	76.0	10894	4	AA836611	AA836611 Human car
158	15.2	76.0	1899	10	ADIS7348	Adi57348 Human pyu	c 231	15.2	76.0	10894	4	AA836611	AA836611 Human car
159	15.2	76.0	1899	10	ABT17040	Abt17040 Human Mf2	c 232	15.2	76.0	10894	4	AA836611	AA836611 Human car
160	15.2	76.0	1899	10	ABT17040	Abt17040 Human Mf2	c 233	15.2	76.0	10894	4	AA836611	AA836611 Human car
161	15.2	76.0	1899	12	ADU62737	Adj62737 Human cdn	c 234	15.2	76.0	10894	10	AA836612	AA836612 Human car
162	15.2	76.0	1899	12	ADU62737	Adj62737 Human cdn	c 235	15.2	76.0	10897	4	AA836612	AA836612 Human car
163	15.2	76.0	1899	12	ADNO3687	Adno3687 Antipsori	c 236	15.2	76.0	10897	4	AA836612	AA836612 Human car
164	15.2	76.0	1899	12	ADNO3687	Adno3687 Antipsori	c 237	15.2	76.0	10897	4	AA836612	AA836612 Human car
165	15.2	76.0	1899	12	ADOI9113	Adoi9113 Human pro	c 238	15.2	76.0	10897	4	AA836612	AA836612 Human car
166	15.2	76.0	1899	12	ADOI9113	Adoi9113 Human pro	c 239	15.2	76.0	10897	4	AA836612	AA836612 Human car
167	15.2	76.0	1920	6	AA29931	AA29931 Human cal	c 240	15.2	76.0	10897	4	AA836612	AA836612 Human car

241	15.2	76.0	10907	8	ABX60591	Abx60591 CDNA	enco	C 314	14.8	74.0	3934	6	ABK54210	Abk54210 Novel pan
C 242	15.2	76.0	10907	8	ABX60591	Abx60591 CDNA	enco	C 315	14.8	74.0	11171	12	ADM860035	Adm860035 Spitzmyct
C 243	15.2	76.0	10907	10	ADBE47306	ADBE47306 Human car		C 316	14.8	74.0	11171	12	ADM860035	Adm860035 Spitzmyct
C 244	15.2	76.0	10907	10	ADBE47306	ADBE47306 Human car		C 317	14.8	74.0	11171	12	ADN97551	Adn97551 S ambofac
C 245	15.2	76.0	10907	12	ADJ31341	ADJ31341 Human mus		C 318	14.8	74.0	11171	12	ADN97551	Adn97551 S ambofac
C 246	15.2	76.0	10907	12	ADJ31341	ADJ31341 Human mus		C 319	14.8	74.0	23241	5	AAFP97871	AAFP97871 Human neu
C 247	15.2	76.0	12718	4	AAK66346	AAK66346 Human imm		C 320	14.8	74.0	23241	5	AAFP97871	AAFP97871 Human neu
C 248	15.2	76.0	12718	4	AAK66346	AAK66346 Human imm		C 321	14.8	74.0	23241	5	AAFP97870	AAFP97870 Human neu
C 249	15.2	76.0	12718	4	AAJ37607	AAJ37607 Human mus		C 322	14.8	74.0	23241	5	AAFP97870	AAFP97870 Human neu
C 250	15.2	76.0	12718	4	AAJ37607	AAJ37607 Human mus		C 323	14.8	74.0	42291	9	ACCS58253	ACCS58253 Clorobloc
C 251	15.2	76.0	12718	8	ABX60595	Abx60595 CDNA	enco	C 324	14.8	74.0	42291	9	ACCS58253	ACCS58253 Clorobloc
C 252	15.2	76.0	12718	8	ABX60595	Abx60595 CDNA	enco	C 325	14.8	74.0	63209	4	ABL19214	ABL19214 Drosophil
C 253	15.2	76.0	12718	12	ADJ31345	ADJ31345 Human mus		C 326	14.8	74.0	63209	4	ABL19214	ABL19214 Drosophil
C 254	15.2	76.0	12718	12	ADJ31345	ADJ31345 Human mus		C 327	14.8	74.0	87350	2	AAAX83003	AAAX83003 Human WRN
C 255	15.2	76.0	51855	10	ACFO4818	ACFO4818 Melithiaz		C 328	14.8	74.0	87350	2	AAAX83003	AAAX83003 Human WRN
C 256	15.2	76.0	51855	10	ACFO4818	ACFO4818 Melithiaz		C 329	14.8	74.0	87543	7	ADJ84092	ADJ84092 Human WRN
C 257	15.2	76.0	67191	9	ADA02903	Ada02903 Mouse Mcm		C 330	14.8	74.0	87543	7	ADJ84092	ADJ84092 Human WRN
C 258	15.2	76.0	67191	9	ADA02903	Ada02903 Mouse Mcm		C 331	14.4	72.0	187	3	AAFI5330	AAFI5330 Trichoder
C 259	15.2	76.0	67191	10	ADB72641	ADB72641 Mouse Mcm		C 332	14.4	72.0	187	3	AAFI5330	AAFI5330 Trichoder
C 260	15.2	76.0	67191	10	ADB72641	ADB72641 Mouse Mcm		C 333	14.4	72.0	370	4	AAAH35407	AAAH35407 Human col
C 261	15.2	76.0	67191	11	ADL95495	ADL95495 Murine MC		C 334	14.4	72.0	370	4	AAAH35407	AAAH35407 Human col
C 262	15.2	76.0	67191	11	ADL95495	ADL95495 Murine MC		C 335	14.4	72.0	473	12	ADJ42779	ADJ42779 Plant CDN
C 263	15.2	76.0	67191	12	ADM74498	Adm74498 Murine ca		C 336	14.4	72.0	473	12	ADJ42779	ADJ42779 Plant CDN
C 264	15.2	76.0	67191	12	ADM74498	Adm74498 Murine ca		C 337	14.4	72.0	697	12	ADJ42781	ADJ42781 Plant CDN
C 265	15.2	76.0	110000	4	AAI96682_02	Continuati	(3 of	C 338	14.4	72.0	697	12	ADJ42781	ADJ42781 Plant CDN
C 266	15.2	76.0	110000	4	AAI96682_02	Continuati	(3 of	C 339	14.4	72.0	1150	12	ADJ42780	ADJ42780 Plant CDN
C 267	15.2	76.0	110000	4	AAI96682_02	Continuati	(3 of	C 340	14.4	72.0	1150	12	ADJ42780	ADJ42780 Plant CDN
C 268	15.2	76.0	110000	4	AAI96683_02	Continuati	(3 of	C 341	14.4	72.0	2002	10	ADG1058	

387	14.2	71.0	566	5	ABAI1734	Abai1734	Human	ner	C 460	14.2	71.0	1326	8	ABT19177	Abt19177	Aspergill
388	14.2	71.0	566	3	ABAI1734	Abai1734	Human	ner	C 461	14.2	71.0	1326	8	ABT20997	Abt20997	Aspergill
389	14.2	71.0	568	3	AAO1081	Aao1081	Human	sec	C 462	14.2	71.0	1326	8	ABT20997	Abt20997	Aspergill
390	14.2	71.0	568	3	AAO1081	Aao1081	Human	sec	C 463	14.2	71.0	1344	4	AAFG1030	Aafg1030	P. putida
391	14.2	71.0	579	4	AAE72100	Aae72100	Corynebac		C 464	14.2	71.0	1344	4	AAFG1030	Aafg1030	P. putida
392	14.2	71.0	579	4	AAE72100	Aae72100	Corynebac		C 465	14.2	71.0	1349	2	AAAT5764	Aat5764	Mouise non
393	14.2	71.0	579	4	AAE72106	Aae72106	Corynebac		C 466	14.2	71.0	1349	2	AAAT5764	Aat5764	Mouise non
394	14.2	71.0	579	4	AAE72106	Aae72106	Corynebac		C 467	14.2	71.0	1488	11	ABD06334	Abd06334	Pseudomon
395	14.2	71.0	579	4	AAE72092	Aae72092	Corynebac		C 468	14.2	71.0	1488	11	ABD06334	Abd06334	Pseudomon
396	14.2	71.0	580	4	AAE72092	Aae72092	Corynebac		C 469	14.2	71.0	1611	8	ACA42277	Aca42277	Prokaryot
397	14.2	71.0	580	4	AAE72092	Aae72092	Corynebac		C 470	14.2	71.0	1611	8	ACA42277	Aca42277	Prokaryot
398	14.2	71.0	580	4	AAE72092	Aae72092	Corynebac		C 471	14.2	71.0	1616	4	AAI66429	Aai66429	Human chr
399	14.2	71.0	589	12	ADJ42130	Adj42130	Plant	CDN	C 472	14.2	71.0	1616	4	AAI66429	Aai66429	Human chr
400	14.2	71.0	589	12	ADJ42130	Adj42130	Plant	CDN	C 473	14.2	71.0	1668	11	ABD15101	Abd15101	Pseudomon
401	14.2	71.0	590	5	ABV59325	Abv59325	Human	pro	C 474	14.2	71.0	1668	11	ABD15101	Abd15101	Pseudomon
402	14.2	71.0	590	5	ABV59325	Abv59325	Human	pro	C 475	14.2	71.0	1726	10	ADD47080	Add47080	Human gen
403	14.2	71.0	591	10	AD122499	Ad122499	Rat	liver	C 476	14.2	71.0	1726	10	ADD47080	Add47080	Human gen
404	14.2	71.0	591	10	AD122499	Ad122499	Rat	liver	C 477	14.2	71.0	1740	4	AAFG1640	Aafg1640	Human
405	14.2	71.0	600	12	ACH69173	Ach69173	Human	gen	C 478	14.2	71.0	1740	4	AAFG1640	Aafg1640	Human
406	14.2	71.0	600	12	ACH69173	Ach69173	Human	gen	C 479	14.2	71.0	1803	4	AAFG3447	Aafg3447	Human
407	14.2	71.0	606	12	ACH69173	Ach69173	Human	gen	C 480	14.2	71.0	1803	4	AAFG3447	Aafg3447	Human
408	14.2	71.0	636	3	AAFI3652	Aafi3652	Aspergill		C 481	14.2	71.0	1803	4	AAFG3447	Aafg3447	Human
409	14.2	71.0	636	3	AAFI3652	Aafi3652	Aspergill		C 482	14.2	71.0	1869	2	AAZ34146	Aaz34146	Human
410	14.2	71.0	636	3	AAFI3652	Aafi3652	Aspergill		C 483	14.2	71.0	1869	2	AAZ34146	Aaz34146	Human
411	14.2	71.0	636	3	AAFI3652	Aafi3652	Aspergill		C 484	14.2	71.0	1869	3	AACT8532	Aac78532	Human
412	14.2	71.0	693	11	ABD14590	Abd14590	Pseudomon		C 485	14.2	71.0	1869	3	AACT8532	Aac78532	Human
413	14.2	71.0	693	11	ABD14590	Abd14590	Pseudomon		C 486	14.2	71.0	1869	3	AACT8532	Aac78532	Human
414	14.2	71.0	697	6	ABZ78094	Abz78094	Human	bre	C 487	14.2	71.0	1869	3	AACT8532	Aac78532	Human
415	14.2	71.0	697	6	ABZ78094	Abz78094	Human	bre	C 488	14.2	71.0	1869	3	AACT8532	Aac78532	Human
416	14.2	71.0	697	6	ABZ78094	Abz78094	Human	bre	C 489	14.2	71.0	1869	3	AACT8532	Aac78532	Human
417	14.2	71.0	726	11	ABD06393	Abd06393	Pseudomon		C 490	14.2	71.0	1869	3	AACT8532	Aac78532	Human
418	14.2	71.0	726	11	ABD06393	Abd06393	Pseudomon		C 491	14.2	71.0	1869	3	AACT8532	Aac78532	Human
419	14.2	71.0	749	4	AAK74242	Aak74242	Human	imm	C 492	14.2	71.0	1869	3	AACT8532	Aac78532	Human
420	14.2	71.0	749	4	AAK74242	Aak74242	Human	imm	C 493	14.2	71.0	1869	3	AACT8532	Aac78532	Human
421	14.2	71.0	749	4	AAK74242	Aak74242	Human	imm	C 494	14.2	71.0	1869	3	AACT8532	Aac78532	Human
422	14.2	71.0	749	4	AAK74242	Aak74242	Human	imm	C 495	14.2	71.0	1869	3	AACT8532	Aac78532	Human
423	14.2	71.0	750	4	AAK74243	Aak74243	Human	imm	C 496	14.2	71.0	1869	3	AACT8532	Aac78532	Human
424	14.2	71.0	750	4	AAK74243	Aak74243	Human	imm	C 497	14.2	71.0	1869	3	AACT8532	Aac78532	Human
425	14.2	71.0	759	11	ABD15110	Abd15110	Pseudomon		C 498	14.2	71.0	1869	3	AACT8532	Aac78532	Human
426	14.2	71.0	759	11	ABD15110	Abd15110	Pseudomon		C 499	14.2	71.0	1869	3	AACT8532	Aac78532	Human
427	14.2	71.0	798	11	ABD14825	Abd14825	Pseudomon		C 500	14.2	71.0	1869	3	AACT8532	Aac78532	Human
428	14.2	71.0	798	11	ABD14825	Abd14825	Pseudomon									
429	14.2	71.0	807	6	ABN85857	Abn85857	Nucleotid									
430	14.2	71.0	807	6	ABN85857	Abn85857	Nucleotid									
431	14.2	71.0	816	11	ABD06347	Abd06347	Pseudomon									
432	14.2	71.0	816	11	ABD06347	Abd06347	Pseudomon									
433	14.2	71.0	816	11	ABD06347	Abd06347	Pseudomon									
434	14.2	71.0	816	11	ABD06347	Abd06347	Pseudomon									
435	14.2	71.0	924	8	ACA4058	Aca4058	Prokaryot									
436	14.2	71.0	924	8	ACA4058	Aca4058	Prokaryot									
437	14.2	71.0	966	11	ABD14833	Abd14833	Pseudomon									
438	14.2	71.0	966	11	ABD14833	Abd14833	Pseudomon									
439	14.2	71.0	1008	8	ACA45287	Aca45287	Prokaryot									
440	14.2	71.0	1008	8	ACA45287	Aca45287	Prokaryot									
441	14.2	71.0	1092	6	ABL90801	Ab190801	Human	pol								
442	14.2	71.0	1092	6	ABL90801	Ab190801	Human	pol								
443	14.2	71.0	1093	5	ADM19375	Adm19375	Novel	hum								
444	14.2	71.0	1093	5	ADM19375	Adm19375	Novel	hum								
445	14.2	71.0	1125	11	ABD12598	Abd12598	Pseudomon									
446	14.2	71.0	1125	11	ABD12598	Abd12598	Pseudomon									
447	14.2	71.0	1134	11	ABD14157	Abd14157	Pseudomon									
448	14.2	71.0	1134	11	ABD14157	Abd14157	Pseudomon									
449	14.2	71.0	1200	6	AAJ55969	Aaj55969	Corynebac									
450	14.2	71.0	1200	6	AAJ55969	Aaj55969	Corynebac									
451	14.2	71.0	1254	11	ABD12038	Abd12038	Pseudomon									
452	14.2	71.0	1254	11	ABD12038	Abd12038	Pseudomon									
453	14.2	71.0	1287	11	ADL27124	Adl27124	Human	cod								
454	14.2	71.0	1287	11	ADL27124	Adl27124	Human	cod								
455	14.2	71.0	1289	3	AACT5770	Aac75770	Human	ORF								
456	14.2	71.0	1289	3	AACT5770	Aac75770	Human	ORF								
457	14.2	71.0	1308	11	ABD14010	Abd14010	Pseudomon									
458	14.2	71.0	1308	11	ABD14010	Abd14010	Pseudomon									
459	14.2	71.0	1326	8	ABT19177	Abt19177	Aspergill									

## ALIGNMENTS

RESULT 1	AAQ88279	standard; cDNA; 20 BP.
AAQ88279	AAQ88279	
AC	AAQ88279	
XX		
DT	25-MAR-2003	(revised)
DT	06-DEC-1995	(first entry)
XX		
DE	p53 protein binding site sequence used to treat neurodegeneration.	
XX		
KW	Antisense; recombinant virus; vector; p53; binding site; Huntington's;	
KW	neuronal degeneration; Parkinson's; Alzheimer's;	
KW	amyotrophic lateral sclerosis; ds.	
XX		
OS	Synthetic.	
XX		
PN	WO950916-A1.	
XX		
PD	13-APR-1995.	
XX		
PF	29-SEP-1994;	94WO-FR001142.
XX		
PR	04-OCT-1993;	93FR-00011774.
XX		
PA	(RHON ) RHON POULENC RORER SA.	
XX		

PI Mallet J, Revah F, Stutzmann J;  
 XX WPI, 1995-155257/20.  
 XX  
 PT Treating and preventing neurodegenerative diseases - with cpds. that  
 PT inhibit p53 activity, partic. an antisense sequence or viral vector.  
 XX  
 PS Claim 12, Page 11, 21pp; French.  
 XX  
 CC The sequence of the p53 binding site sequence used in a composition to  
 CC treat neuronal degeneration. The sequence may be used as an  
 CC oligonucleotide used either solely or as part of a recombinant viral  
 CC vector to inhibit the activity of the p53 protein. The oligonucleotide  
 CC may be used in conjunction, either as naked DNA or on the same  
 CC recombinant viral vector, with a p53 antisense sequence which inhibits  
 CC expression of the p53 gene (AA088280). The viral vectors may also encode  
 CC a mutated form of p53 able to antagonise p53 protein activity. These  
 CC oligonucleotides and recombinant viruses can be used in compositions to  
 CC treat and prevent neuronal degeneration, Huntington's, Parkinson's or  
 CC Alzheimer's diseases or amyotrophic lateral sclerosis. (Updated on 25-MAR  
 CC -2003 to correct PN field.)  
 CC  
 SQ Sequence 20 BP, 3 A, 7 C, 7 G, 3 T, 0 U, 0 Other;  
 Query Match 100.0%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 1 GGACATGCCCGGCGCATGTCC 20  
 DB  
 RESULT 2  
 AA088279/c  
 ID AA088279 standard; cDNA; 20 BP.  
 XX  
 AC AA088279;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 06-DEC-1995 (first entry)  
 XX  
 DE p53 protein binding site sequence used to treat neurodegeneration.  
 XX  
 KM Antisense; recombinant virus; vector; p53; binding site; Huntington's;  
 KM neuronal degeneration; Parkinson's; Alzheimer's;  
 KM amyotrophic lateral sclerosis; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9509916-A1.  
 XX  
 PD 13-APR-1995.  
 XX  
 PF 29-SEP-1994; 94WO-FR001142.  
 XX  
 PR 04-OCT-1993; 93FR-00011774.  
 XX  
 PA (RHON) RHON POULENC RORER SA.  
 XX  
 PI Mallet J, Revah F, Stutzmann J;  
 XX  
 DR WPI, 1995-155257/20.  
 XX  
 PT Treating and preventing neurodegenerative diseases - with cpds. that  
 PT inhibit p53 activity, partic. an antisense sequence or viral vector.  
 XX  
 PS Claim 12, Page 11, 21pp; French.  
 XX  
 CC The sequence of the p53 binding site sequence used in a composition to  
 CC treat neuronal degeneration. The sequence may be used as an  
 CC oligonucleotide used either solely or as part of a recombinant viral  
 CC vector to inhibit the activity of the p53 protein. The oligonucleotide

CC may be used in conjunction, either as naked DNA or on the same  
 CC recombinant viral vector, with a p53 antisense sequence which inhibits  
 CC expression of the p53 gene (AA088280). The viral vectors may also encode  
 CC a mutated form of p53 able to antagonise p53 protein activity. These  
 CC oligonucleotides and recombinant viruses can be used in compositions to  
 CC treat and prevent neuronal degeneration, Huntington's, Parkinson's or  
 CC Alzheimer's diseases or amyotrophic lateral sclerosis. (Updated on 25-MAR  
 CC -2003 to correct PN field.)  
 CC  
 SQ Sequence 20 BP, 3 A, 7 C, 7 G, 3 T, 0 U, 0 Other;  
 Query Match 100.0%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 20 GGACATGCCCGGCGCATGTCC 1  
 DB  
 RESULT 3  
 AAV22590  
 ID AAV22590 standard; DNA; 20 BP.  
 XX  
 AC AAV22590;  
 XX  
 DT 08-JUL-1998 (first entry)  
 DT  
 XX  
 DE p53 consensus binding sequence.  
 XX  
 KM R1 subunit; ribonucleotide reductase; cell proliferation; tumour cell;  
 KM antisense; growth; inhibition; sensitivity; hydroxyurea; R2 subunit;  
 KM chemotherapeutic drug; methotrexate; PALA; treatment; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9805769-A2.  
 XX  
 PD 12-FEB-1998.  
 XX  
 PF 01-AUG-1997; 97WO-CA000540.  
 XX  
 PR 02-AUG-1996; 96US-0023040P.  
 PR 07-MAR-1997; 97US-0039595P.  
 XX  
 PA (GENE-) GENENSENSE TECHNOLOGIES INC.  
 XX  
 PI Wright JA, Young AH;  
 XX  
 DR WPI, 1998-145609/13.  
 XX  
 PT Antisense oligonucleotides to ribonucleotide reductase genes - used to  
 PT modulate tumour growth and inhibit tumour cell proliferation.  
 XX  
 PS Example 2, Page 24; 79pp; English.  
 XX  
 CC The present sequence represents a p53 consensus binding sequence. The  
 CC specification describes antisense oligonucleotides which are targeted  
 CC against the mRNAs of the R1 and R2 subunit sequences of ribonucleotide  
 CC reductase. Aberrant expression of the R2 gene, which encodes the second  
 CC subunit of the ribonucleotide reductase gene, can determine the malignant  
 CC characteristics of cells. Suppression of R2 and R1 gene expression was  
 CC found to reduce transformed properties of tumour cells. The antisense  
 CC oligonucleotides can be used for modulating tumour cell growth, or for  
 CC inhibiting tumour cell proliferation. They can also be used for  
 CC increasing the sensitivity of neoplastic cells to chemotherapeutic drugs  
 CC (especially to hydroxyurea, methotrexate (MTX), and PALA). The antisense  
 CC oligonucleotides may be used to treat proliferative disorders including  
 CC leukaemias, lymphomas, sarcomas, melanomas, various other forms of  
 CC cancer, papillomas, arthrosclerosis, psoriasis, polyphenia, mastocytosis,  
 CC autoimmune diseases, angiogenesis, bacterial infections and viral  
 CC infections (including HIV hepatitis, or herpes infections)

SO Sequence 20 BP, 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
DB 1 GGACATGCCCGGCGCATGTCC 20

#### RESULT 4

AAV22590/c  
ID AAV22590 standard; DNA; 20 BP.

XX AAV22590;

DT 08-JUL-1998 (first entry)

XX p53 consensus binding sequence.

XX R1 subunit; ribonucleotide reductase; cell proliferation; tumour cell;

XX antisense; growth; inhibition; sensitivity; hydroxyurea; R2 subunit;

XX chemotherapeutic drug; methotrexate; PALA; treatment; ss.

XX Synthetic.

XX MO9805769-A2.

XX 12-FEB-1998.

PF 01-AUG-1997; 97WO-CA000540.

PR 02-AUG-1996; 96US-0023040P.

PR 07-MAR-1997; 97US-003959P.

XX (GENE-) GENESENSE TECHNOLOGIES INC.

XX Wright JA, Young AH;

XX WPI, 1998-145609/13.

XX Antisense oligonucleotides to ribonucleotide reductase genes - used to  
modulate tumour growth and inhibit tumour cell proliferation.

XX Example 2; Page 24; 79pp; English.

CC The present sequence represents a p53 consensus binding sequence. The  
CC specification describes antisense oligonucleotides which are targeted  
CC against the mRNAs of the R1 and R2 subunit sequences of ribonucleotide  
CC reductase. Aberrant expression of the R2 gene, which encodes the second  
CC subunit of the ribonucleotide reductase gene, can determine the malignant  
CC characteristics of cells. Suppression of R2 and R1 gene expression was  
CC found to reduce transformed properties of tumour cells. The antisense  
CC oligonucleotides can be used for modulating tumour cell growth, or for  
CC inhibiting tumour cell proliferation. They can also be used for  
CC increasing the sensitivity of neoplastic cells to chemotherapeutic drugs  
CC (especially to hydroxyurea, methotrexate (MTX), and PALA). The antisense  
CC oligonucleotides may be used to treat proliferative disorders including  
CC leukaemias, lymphomas, sarcomas, melanomas, various other forms of  
CC cancer, papillomas, arteriosclerosis, psoriasis, polythemia, mastocytosis,  
CC autoimmune diseases, angiogenesis, bacterial infections and viral  
CC infections (including HIV hepatitis, or herpes infections)

XX Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.6;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
DB 20 GGACATGCCCGGCGCATGTCC 1

RESULT 5  
AAK27873  
ID AAK27873 standard; DNA; 20 BP.

XX AAK27873;

DT 02-JUN-1999 (first entry)

XX Probe for human CSR protein coding sequence.

XX Cellular stress response protein; CSR1; CSR2; CSR3; human; macrophage;

XX scavenger receptor protein; intracellular stress; arteriosclerosis;

XX diabetic circulatory obstruction; microbial infection; probe; ss.

XX Synthetic.

XX Homo sapiens.

XX MO9909159-A1.

XX 25-FEB-1999.

XX 12-AUG-1998; 98WO-JP003602.

XX 13-AUG-1997; 97JP-00233396.

XX 30-JUL-1998; 98JP-00230121.

XX (NISB) JAPAN TOBACCO INC.

XX Nakamura Y, Tokino T;

XX WPI, 1999-161032/15.

XX Scavenger receptor proteins - for treatment and diagnosis of disorders  
involving cell stress.

XX Example 8; Page 159; 175pp; Japanese.

CC This sequence represents a probe for DNA encoding a human cellular stress  
CC response (CSR) protein of the invention. The CSR proteins are macrophage  
CC scavenger receptor proteins. The CSR proteins can be used in the  
CC treatment, gene therapy and diagnosis of diseases in which intracellular  
CC stress is important, such as arteriosclerosis, diabetic circulatory  
CC obstruction, and microbial infection. Expression of the proteins is  
CC induced in vivo in response to intracellular stress, and inhibits cell  
CC death as a result of such stress

XX Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
DB 1 GGACATGCCCGGCGCATGTCC 20

#### RESULT 6

AAK27873/c  
ID AAK27873 standard; DNA; 20 BP.

XX AAK27873;

DT 02-JUN-1999 (first entry)

XX Probe for human CSR protein coding sequence.

XX Cellular stress response protein; CSR1; CSR2; CSR3; human; macrophage;

XX scavenger receptor protein; intracellular stress; arteriosclerosis;  
XX diabetic circulatory obstruction; microbial infection; probe; ss.

OS Synthetic.  
 OS Homo sapiens.  
 PN W09909159-A1.  
 PD 25-FEB-1999.  
 PF 12-AUG-1998; 98WO-JP003602.  
 XX 13-AUG-1997; 97JP-00233396.  
 PR 30-JUL-1998; 98JP-00230121.  
 XX (NIBS ) JAPAN TOBACCO INC.  
 PA Nakamura Y, Tokino T;  
 PI WPI; 1999-181032/15.  
 DR  
 XX  
 PT Scavenger receptor proteins - for treatment and diagnosis of disorders  
 PT involving cell stresses.  
 XX  
 PS Example 8; Page 159; 175pp; Japanese.  
 XX  
 CC This sequence represents a probe for DNA encoding a human cellular stress  
 CC response (CSR) protein of the invention. The CSR proteins are macrophage  
 CC scavenger receptor proteins. The CSR proteins can be used in the  
 CC treatment, gene therapy and diagnosis of diseases in which intracellular  
 CC stress is important, such as arteriosclerosis, diabetic circulatory  
 CC obstruction, and microbial infection. Expression of the proteins is  
 CC induced in vivo in response to intracellular stress, and inhibits cell  
 CC death as a result of such stress  
 CC  
 SQ Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 DB 20 GGACATGCCCGGCATGTCC 1

RESULT 7  
 ID AAA91613 standard; DNA; 20 BP.  
 AC AAA91613;  
 XX  
 DT 20-DEC-2000 (first entry)  
 XX  
 DE p53 consensus binding sequence.  
 XX  
 KW p53 consensus binding sequence; ribonucleotide reductase; R1 protein;  
 KW R2 protein; antisense inhibition; cancer; cytostatic; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN W0200047733-A1.  
 PD 17-AUG-2000.  
 PF 09-FEB-2000; 2000WO-CA000120.  
 PR 11-FEB-1999; 99US-00249730.  
 XX  
 PA (GENE-) GENESENSE TECHNOLOGIES INC.  
 XX  
 PI Wright JA, Young AH;  
 XX WPI; 2000-558216/51.  
 DR  
 XX New antisense oligonucleotide, AS-I-618-20, is useful for inhibiting

PT tumor cell growth.  
 XX  
 XX Example 2; Page 71; 137pp; English.  
 XX

CC The present sequence is a double-stranded p53 consensus binding sequence.  
 CC It was used in an electrophoretic gel mobility shift assay (EMSA) to  
 CC determine the presence of wild type p53 in various mouse cell lines. The  
 CC sequence was end-labelled with radioactive ATP using T4 polynucleotide  
 CC kinase and an excess was incubated with the nuclear lysate from mouse  
 CC cells. This was performed as part of a method for modulating cell  
 CC proliferation using antisense oligonucleotides directed against the R1  
 CC and R2 components of mammalian ribonucleotide reductase. Ribonucleotide  
 CC reductase catalyses the conversion of ribonucleotides to their  
 CC corresponding deoxyribonucleotides and thus plays an important role in  
 CC DNA synthesis and cell proliferation. Regulation of ribonucleotide  
 CC reductase is altered in cultured malignant cells and increased levels of  
 CC R2 protein and R2 mRNA have been found in pre-malignant and malignant  
 CC tissues as compared to normal control tissue samples. The antisense  
 CC sequences are therefore useful for inhibiting tumorigenicity of  
 CC neoplastic cells and inhibiting metastasis of tumour cells. They are also  
 CC useful for increasing sensitivity of neoplastic cells to chemotherapeutic  
 CC drugs, thus allowing chemotherapeutic treatments to be used in patients  
 CC who have become resistant or less sensitive to chemotherapy  
 CC  
 SQ Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 DB 1 GGACATGCCCGGCATGTCC 20

RESULT 8  
 ID AAA91613/c  
 AAA91613 standard; DNA; 20 BP.

AC AAA91613;  
 XX  
 DT 20-DEC-2000 (first entry)  
 XX  
 DE p53 consensus binding sequence.  
 XX  
 KW p53 consensus binding sequence; ribonucleotide reductase; R1 protein;  
 KW R2 protein; antisense inhibition; cancer; cytostatic; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN W0200047733-A1.  
 PD 17-AUG-2000.  
 PF 09-FEB-2000; 2000WO-CA000120.  
 PR 11-FEB-1999; 99US-00249730.  
 XX  
 PA (GENE-) GENESENSE TECHNOLOGIES INC.  
 XX  
 PI Wright JA, Young AH;  
 XX WPI; 2000-558216/51.  
 DR  
 XX New antisense oligonucleotide, AS-I-618-20, is useful for inhibiting  
 PT tumor cell growth.  
 XX  
 PS Example 2; Page 71; 137pp; English.  
 XX  
 CC The present sequence is a double-stranded p53 consensus binding sequence.  
 CC It was used in an electrophoretic gel mobility shift assay (EMSA) to  
 CC determine the presence of wild type p53 in various mouse cell lines. The  
 CC sequence was end-labelled with radioactive ATP using T4 polynucleotide

CC Kinase and an excess was incubated with the nuclear lysate from mouse  
 CC cells. This was performed as part of a method for modulating cell  
 CC proliferation using antisense oligonucleotides directed against the R1  
 CC and R2 components of mammalian ribonucleotide reductase. Ribonucleotide  
 CC reductase catalyzes the conversion of ribonucleotides to their  
 CC corresponding deoxyribonucleotides and thus plays an important role in  
 CC DNA synthesis and cell proliferation. Regulation of ribonucleotide  
 CC reductase is altered in cultured malignant cells and increased levels of  
 CC R2 protein and R2 mRNA have been found in pre-malignant and malignant  
 CC tissues as compared to normal control tissue samples. The antisense  
 CC sequences are therefore useful for inhibiting tumorigenicity of  
 CC neoplastic cells and inhibiting metastasis of tumour cells. They are also  
 CC useful for increasing sensitivity of neoplastic cells to chemotherapeutic  
 CC drugs, thus allowing chemotherapeutic treatments to be used in patients  
 CC who have become resistant or less sensitive to chemotherapy

XX Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCATGTCC 20  
 |||||  
 DB 20 GGACATGCCGGGCATGTCC 1

# RESULT 9

AAD27540  
 ID AAD27540 standard; DNA; 20 BP.

AC AAD27540;

XX 18-Apr-2002 (first entry)

DE p53 protein binding consensus sequence, p53CON.

XX p53 protein; pGL3 luciferase reporter vector; luc+; transcription factor;  
 KM cell cycle control; DNA damage repair; p53CON; apoptosis; ds.

XX Unidentified.

XX WO200196602-A2.

XX 20-DEC-2001.

XX 18-JUN-2001; 2001WO-GB002718.

XX 16-JUN-2000; 2000GB-00014820.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Yang AL, Feasting M;

XX WPI; 2002-130743/17.

XX Determining the p53 status of a sample, useful for assaying for mimetics  
 PT or antagonists of p53, or for the presence of DNA damage, comprises  
 PT determining whether p53 binds to the pGL3 vector in a sample containing a  
 PT pGL3 vector.

XX Disclosure; Page 7; 53pp; English.

XX The patent discloses methods for determining the p53 status of a sample  
 CC which comprise providing a sample containing a pGL3 luciferase reporter  
 CC vector and determining whether p53 binds to the pGL3 vector. p53 is a  
 CC transcription factor that regulates many genes including those associated  
 CC with cell cycle control, apoptosis and DNA damage repair. pGL3 reporter  
 CC vectors contain a modified firefly luciferase cDNA designated luc+. p53  
 CC protein binds to pGL3-basic vector and causes luciferase expression. The  
 CC method is useful for determining the p53 status of a sample. It is also  
 CC useful for assaying for mimetics or antagonists of p53 and for assaying  
 CC for presence of activated p53 and/or DNA damage. The invention also

CC relates to a method of modifying pGL3 vector which involves deletion or  
 CC alteration of a CCCGGG motif of the pGL3 vector and/or deleting or  
 CC altering a sequence within 20 bp sequence 5' or 3' of CCCGGG motif. The  
 CC nucleic acid having a sequence incorporating the CCCGGG motif is useful  
 CC for conferring promoter activity or p53-responsiveness on a nucleic acid  
 CC encoding a polypeptide of interest or in assays for p53 transcriptional  
 CC activity. The present DNA sequence is p53 protein binding consensus  
 CC sequence containing a CCCGGG motif, p53CON

XX Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCATGTCC 20  
 |||||  
 DB 1 GGACATGCCGGGCATGTCC 20

# RESULT 10

AAD27540/c  
 ID AAD27540 standard; DNA; 20 BP.

AC AAD27540;

XX 18-Apr-2002 (first entry)

DE p53 protein binding consensus sequence, p53CON.

XX p53 protein; pGL3 luciferase reporter vector; luc+; transcription factor;  
 KM cell cycle control; DNA damage repair; p53CON; apoptosis; ds.

XX Unidentified.

XX WO200196602-A2.

XX 20-DEC-2001.

XX 18-JUN-2001; 2001WO-GB002718.

XX 16-JUN-2000; 2000GB-00014820.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Yang AL, Feasting M;

XX WPI; 2002-130743/17.

XX Determining the p53 status of a sample, useful for assaying for mimetics  
 PT or antagonists of p53, or for the presence of DNA damage, comprises  
 PT determining whether p53 binds to the pGL3 vector in a sample containing a  
 PT pGL3 vector.

XX Disclosure; Page 7; 53pp; English.

XX The patent discloses methods for determining the p53 status of a sample  
 CC which comprise providing a sample containing a pGL3 luciferase reporter  
 CC vector and determining whether p53 binds to the pGL3 vector. p53 is a  
 CC transcription factor that regulates many genes including those associated  
 CC with cell cycle control, apoptosis and DNA damage repair. pGL3 reporter  
 CC vectors contain a modified firefly luciferase cDNA designated luc+. p53  
 CC protein binds to pGL3-basic vector and causes luciferase expression. The  
 CC method is useful for determining the p53 status of a sample. It is also  
 CC useful for assaying for mimetics or antagonists of p53 and for assaying  
 CC for presence of activated p53 and/or DNA damage. The invention also  
 CC relates to a method of modifying pGL3 vector which involves deletion or  
 CC alteration of a CCCGGG motif of the pGL3 vector and/or deleting or  
 CC altering a sequence within 20 bp sequence 5' or 3' of CCCGGG motif. The  
 CC nucleic acid having a sequence incorporating the CCCGGG motif is useful  
 CC for conferring promoter activity or p53-responsiveness on a nucleic acid  
 CC encoding a polypeptide of interest or in assays for p53 transcriptional  
 CC activity. The present DNA sequence is p53 protein binding consensus



CC sequence containing a CCCGGG motif, p53CON  
XX  
SQ Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGACATGCCCGGGCATGTCC 20  
DB 20 GGACATGCCCGGGCATGTCC 1  
RESULT 11  
ABK98014  
ID ABK98014 standard; DNA; 24 BP.  
AC ABK98014;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Cell-TRAP method associated Oct oligonucleotide.  
XX  
KM Transcription factor; transcription factor-responsive element; ds; TFRB;  
KM transcription activation; Cell-TRAP.  
OS Synthetic.  
XX  
PN WO200252039-A2.  
XX  
PD 04-JUL-2002.  
XX  
PF 21-DEC-2001; 2001WO-CA001861.  
XX  
PR 27-DEC-2000; 2000CA-02327581.  
XX  
PA (GENE-) GENEKA BIOTECHNOLOGY INC.  
PI Blais Y, Rousseau P, Leblanc B, Camato RN;  
XX  
DR WPI; 2002-575388/61.  
XX  
PT A Cell-TRAP method, useful for producing or validating therapeutic  
PT compounds, by employing a recombinant cell-based library that carry  
PT constructs driven by a minimal promoter and a transcription factor-  
PT responsive element.  
XX  
PS Disclosure; Page 25; 44pp; English.  
XX  
CC This invention relates to a cell-TRAP method for selecting and producing  
CC a therapeutic compound which is presumed selective for, one or a  
CC restricted set of given transcriptional pathways and cell types by  
CC employing a recombinant cell-based library that carries a construct  
CC comprising a reporter gene driven by a minimal promoter and a  
CC transcription factor-responsive element (TFRB). The invention also  
CC comprises a method for validating a putative compound as a selective  
CC therapeutic compound towards a transcription factor response element. The  
CC method of the invention is useful for determining the transcriptional  
CC activation pathways used by any compound that is biologically active in a  
CC cell. This method allows a global view of gene transcription activation  
CC in response to diverse stimuli in multiple environments and is a  
CC significant improvement over case-by-case approaches, which would be  
CC limited to certain aspects of gene activation. It permits to save on  
CC clinical trials by screening properly the compounds that would have a  
CC lesser probability of providing undesirable, even severe side effects.  
CC The present sequence represents a double stranded oligonucleotide probe  
CC recognised by a specific transcription factor which is used in the method  
CC of the invention  
XX  
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 6; Length 24;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGACATGCCCGGGCATGTCC 20  
DB 5 GGACATGCCCGGGCATGTCC 24  
RESULT 12  
ABK98014/C  
ID ABK98014 standard; DNA; 24 BP.  
AC ABK98014;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Cell-TRAP method associated Oct oligonucleotide.  
XX  
KM Transcription factor; transcription factor-responsive element; ds; TFRB;  
KM transcription activation; Cell-TRAP.  
OS Synthetic.  
XX  
PN WO200252039-A2.  
XX  
PD 04-JUL-2002.  
XX  
PF 21-DEC-2001; 2001WO-CA001861.  
XX  
PR 27-DEC-2000; 2000CA-02327581.  
XX  
PA (GENE-) GENEKA BIOTECHNOLOGY INC.  
PI Blais Y, Rousseau P, Leblanc B, Camato RN;  
XX  
DR WPI; 2002-575388/61.  
XX  
PT A Cell-TRAP method, useful for producing or validating therapeutic  
PT compounds, by employing a recombinant cell-based library that carry  
PT constructs driven by a minimal promoter and a transcription factor-  
PT responsive element.  
XX  
PS Disclosure; Page 25; 44pp; English.  
XX  
CC This invention relates to a cell-TRAP method for selecting and producing  
CC a therapeutic compound which is presumed selective for, one or a  
CC restricted set of given transcriptional pathways and cell types by  
CC employing a recombinant cell-based library that carries a construct  
CC comprising a reporter gene driven by a minimal promoter and a  
CC transcription factor-responsive element (TFRB). The invention also  
CC comprises a method for validating a putative compound as a selective  
CC therapeutic compound towards a transcription factor response element. The  
CC method of the invention is useful for determining the transcriptional  
CC activation pathways used by any compound that is biologically active in a  
CC cell. This method allows a global view of gene transcription activation  
CC in response to diverse stimuli in multiple environments and is a  
CC significant improvement over case-by-case approaches, which would be  
CC limited to certain aspects of gene activation. It permits to save on  
CC clinical trials by screening properly the compounds that would have a  
CC lesser probability of providing undesirable, even severe side effects.  
CC The present sequence represents a double stranded oligonucleotide probe  
CC recognised by a specific transcription factor which is used in the method  
CC of the invention  
XX  
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 6; Length 24;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGACATGCCCGGGCATGTCC 20  
DB 24 GGACATGCCCGGGCATGTCC 5

RESULT 13  
ABK98249  
ID ABK98249 standard; DNA; 24 BP.  
XX  
AC ABK98249;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Nucleic acid binding element (NABE) probe #19.  
XX  
KW Nucleic acid binding element; NABE; nucleic acid binding factor; NABF;  
KW probe; ss; nucleic acid binding protein; DNA microarray;  
KW viral integration.  
XX  
OS Synthetic.  
XX  
PN WO200252037-A2.  
XX  
PD 04-JUL-2002.  
XX  
PF 14-DEC-2001; 2001WO-CA001794.  
XX  
PR 27-DEC-2000; 2000CA-02327561.  
XX  
PA (GENE-) GENEKA BIOTECHNOLOGY INC.  
XX  
PI Larose A, Rousseau P, Leblanc B, Camato R;  
XX  
DR WPI; 2002-575386/61.  
XX  
PT Screening for nucleic acid binding elements (NABE) or factors (NABF),  
PT useful for analyzing DNA-protein interactions, by identifying NABE-NABF  
PT complexes using an electromobility shift assay, then identifying NABE  
PT using NABE probes.  
XX  
PS Example 4; Page 22; 37bp; English.  
XX  
CC The invention relates to a method for screening nucleic acid binding  
CC elements (NABEs) that are differentially active in modified cells,  
CC comprising identifying complexes between NABEs and nucleic acid binding  
CC factors (NABFs) in order to separate and obtain NABEs that bind to NABFs,  
CC followed by a hybridisation assay to identify marked NABEs using NABE  
CC probes. The method is useful in rapid screening and large-scale  
CC identification or analyses of NABEs or NABFs that are differentially  
CC active in cells that have undergone particular treatments or that are  
CC phenotypically different, using DNA microarray technology. These nucleic  
CC acid binding proteins are involved in a variety of cellular processes  
CC ranging from transcription and replication to recombination and viral  
CC integration. The method is also useful in providing simultaneous analysis  
CC of the effect of a given compound on multiple DNA-protein interactions,  
CC or for the analysis of compounds that are nucleic acid binding factor  
CC analogues, and that selectively bind cis-acting nucleic acids. Sequences  
CC ABK98231-ABK98258 represent NABE probes used to make DNA microarrays  
XX  
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;  
XX  
QY Query Match 100.0%; Score 20; DB 6; Length 24;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 GGACATGCCCGGCGCATGTCC 20  
5 GGACATGCCCGGCGCATGTCC 24

XX  
DE Nucleic acid binding element (NABE) probe #19.  
XX  
KW Nucleic acid binding element; NABE; nucleic acid binding factor; NABF;  
KW probe; ss; nucleic acid binding protein; DNA microarray;  
KW viral integration.  
XX  
OS Synthetic.  
XX  
PN WO200252037-A2.  
XX  
PD 04-JUL-2002.  
XX  
PF 14-DEC-2001; 2001WO-CA001794.  
XX  
PR 27-DEC-2000; 2000CA-02327561.  
XX  
PA (GENE-) GENEKA BIOTECHNOLOGY INC.  
XX  
PI Larose A, Rousseau P, Leblanc B, Camato R;  
XX  
DR WPI; 2002-575386/61.  
XX  
PT Screening for nucleic acid binding elements (NABE) or factors (NABF),  
PT useful for analyzing DNA-protein interactions, by identifying NABE-NABF  
PT complexes using an electromobility shift assay, then identifying NABE  
PT using NABE probes.  
XX  
PS Example 4; Page 22; 37bp; English.  
XX  
CC The invention relates to a method for screening nucleic acid binding  
CC elements (NABEs) that are differentially active in modified cells,  
CC comprising identifying complexes between NABEs and nucleic acid binding  
CC factors (NABFs) in order to separate and obtain NABEs that bind to NABFs,  
CC followed by a hybridisation assay to identify marked NABEs using NABE  
CC probes. The method is useful in rapid screening and large-scale  
CC identification or analyses of NABEs or NABFs that are differentially  
CC active in cells that have undergone particular treatments or that are  
CC phenotypically different, using DNA microarray technology. These nucleic  
CC acid binding proteins are involved in a variety of cellular processes  
CC ranging from transcription and replication to recombination and viral  
CC integration. The method is also useful in providing simultaneous analysis  
CC of the effect of a given compound on multiple DNA-protein interactions,  
CC or for the analysis of compounds that are nucleic acid binding factor  
CC analogues, and that selectively bind cis-acting nucleic acids. Sequences  
CC ABK98231-ABK98258 represent NABE probes used to make DNA microarrays  
XX  
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;  
XX  
QY Query Match 100.0%; Score 20; DB 6; Length 24;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 GGACATGCCCGGCGCATGTCC 20  
24 GGACATGCCCGGCGCATGTCC 5

RESULT 15  
ADP26897  
ID ADP26897 standard; DNA; 24 BP.  
XX  
AC ADP26897;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Transcription factor activation analysis-related oligonucleotide Segrid1.  
XX  
KW activation; transcription factor; labelling; ss.  
XX  
OS Unidentified.  
XX  
PN JP2004159609-A.

```

XX 10-JUN-2004.
PD 15-NOV-2002; 2002JP-00332024.
XX
PF 15-NOV-2002; 2002JP-00332024.
XX
PR 15-NOV-2002; 2002JP-00332024.
XX
PA (FUJI/) FUJIMORI F.
PA (IKED/) IKEDA H.
PA (INIT-) INITIUM KK.
XX
DR WPI; 2004-424491/40.
XX
PT Analyzing grade of activation of transcription factor in test substance,
PT involves hybridizing probe labeled with specific substance, to test
PT substance, detecting labeled probe and recovering probe coupled with
PT transcription factor.
XX
PS Example 1; Page 16; 12pp; Japanese.
XX
CC This invention relates to a novel method of analyzing the grade of
CC activation of transcription factor in a test substance, which involves
CC hybridizing a labelled probe to a test substance, where the probe is
CC labelled with different labelling substance specific for each test
CC substance, detecting the labelled probe, and recovering the probe coupled
CC with transcription factor on a single plate. The invention is useful for
CC analyzing the grade of activation of transcription factor in a test
CC substance efficiently and with high sensitivity. The present sequence is
CC that of an oligonucleotide which was used in the exemplification of the
CC invention.
XX
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GGACATGCCCGGCATGTCC 20
Db 1 GGACATGCCCGGCATGTCC 20
XX
RESULT 16
ADP26897/c
ID ADP26897 standard; DNA; 24 BP.
XX
AC ADP26897;
XX
DT 26-AUG-2004 (first entry)
XX
DE Transcription factor activation analysis-related oligonucleotide Segid11.
XX
KW activation; transcription factor; labelling; ss.
XX
OS Unidentified.
XX
PN JP2004159609-A.
XX
PD 10-JUN-2004.
XX
PF 15-NOV-2002; 2002JP-00332024.
XX
PR 15-NOV-2002; 2002JP-00332024.
XX
PA (FUJI/) FUJIMORI F.
PA (IKED/) IKEDA H.
PA (INIT-) INITIUM KK.
XX
DR WPI; 2004-424491/40.
XX
PT Analyzing grade of activation of transcription factor in test substance,
PT involves hybridizing probe labeled with specific substance, to test
PT substance, detecting labeled probe and recovering probe coupled with

```

```

PT transcription factor.
XX
XX Example 1; Page 16; 12pp; Japanese.
XX
CC This invention relates to a novel method of analyzing the grade of
CC activation of transcription factor in a test substance, which involves
CC hybridizing a labelled probe to a test substance, where the probe is
CC labelled with different labelling substance specific for each test
CC substance, detecting the labelled probe, and recovering the probe coupled
CC with transcription factor on a single plate. The invention is useful for
CC analyzing the grade of activation of transcription factor in a test
CC substance efficiently and with high sensitivity. The present sequence is
CC that of an oligonucleotide which was used in the exemplification of the
CC invention.
XX
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GGACATGCCCGGCATGTCC 20
Db 20 GGACATGCCCGGCATGTCC 1
XX
RESULT 17
ADP26896
ID ADP26896 standard; DNA; 24 BP.
XX
AC ADP26896;
XX
DT 26-AUG-2004 (first entry)
XX
DE Transcription factor activation analysis-related oligonucleotide Segid10.
XX
KW activation; transcription factor; labelling; ss.
XX
OS Unidentified.
XX
PN JP2004159609-A.
XX
PD 10-JUN-2004.
XX
PF 15-NOV-2002; 2002JP-00332024.
XX
PR 15-NOV-2002; 2002JP-00332024.
XX
PA (FUJI/) FUJIMORI F.
PA (IKED/) IKEDA H.
PA (INIT-) INITIUM KK.
XX
DR WPI; 2004-424491/40.
XX
PT Analyzing grade of activation of transcription factor in test substance,
PT involves hybridizing probe labeled with specific substance, to test
PT substance, detecting labeled probe and recovering probe coupled with
PT transcription factor.
XX
PS Example 1; Page 16; 12pp; Japanese.
XX
CC This invention relates to a novel method of analyzing the grade of
CC activation of transcription factor in a test substance, which involves
CC hybridizing a labelled probe to a test substance, where the probe is
CC labelled with different labelling substance specific for each test
CC substance, detecting the labelled probe, and recovering the probe coupled
CC with transcription factor on a single plate. The invention is useful for
CC analyzing the grade of activation of transcription factor in a test
CC substance efficiently and with high sensitivity. The present sequence is
CC that of an oligonucleotide which was used in the exemplification of the
CC invention.
XX
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 20; DB 12; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 DB 5 GGACATGCCCGGCGCATGTCC 24

RESULT 18  
 ADP26896/c  
 ID ADP26896 standard; DNA; 24 BP.  
 XX  
 AC ADP26896;  
 XX  
 XX 26-AUG-2004 (first entry)  
 XX  
 DE Transcription factor activation analysis-related oligonucleotide SeqID10.  
 XX  
 KM activation; transcription factor; labelling; ss.  
 XX  
 OS Unidentified.  
 XX  
 PN JP2004159609-A.  
 XX  
 PD 10-JUN-2004.  
 XX  
 PF 15-NOV-2002; 2002JP-00332024.  
 XX  
 PR 15-NOV-2002; 2002JP-00332024.  
 XX  
 PA (FUJII/) FUJIMORI F.  
 PA (IKEDA/) IKEDA H.  
 PA (INIT-) INITIUM KK.  
 XX  
 DR WPI; 2004-424491/40.  
 XX  
 PT Analyzing grade of activation of transcription factor in test substance,  
 PT involves hybridizing probe labeled with specific substance, to test  
 PT substance, detecting labeled probe and recovering probe coupled with  
 PT transcription factor.  
 XX  
 PS Example 1; Page 16; 12pp; Japanese.  
 XX  
 CC This invention relates to a novel method of analysing the grade of  
 CC activation of transcription factor in a test substance, which involves  
 CC hybridizing a labelled probe to a test substance, where the probe is  
 CC labelled with different labelling substance specific for each test  
 CC substance, detecting the labelled probe, and recovering the probe coupled  
 CC with transcription factor on a single plate. The invention is useful for  
 CC analysing the grade of activation of transcription factor in a test  
 CC substance efficiently and with high sensitivity. The present sequence is  
 CC that of an oligonucleotide which was used in the exemplification of the  
 CC invention.  
 CC  
 SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 DB 24 GGACATGCCCGGCGCATGTCC 5

RESULT 19  
 ID AAA61639 standard; DNA; 26 BP.  
 XX  
 AC AAA61639;  
 XX

DT 23-OCT-2000 (first entry)  
 XX  
 DE p53 binding site oligonucleotide, SEQ ID NO:40.  
 XX  
 KM p53 binding site; ubiquitin conjugating enzyme;  
 KM ubiquitin mediated proteolysis; human; cellular protein half life;  
 KM ubiquitin mediated proteolysis; p53; cyclin; cell cycle regulator;  
 KM myc deregulation; human papillomavirus; HPV-18 B6 protein;  
 KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;  
 KM psoriasis; connective tissue disorder; wound healing; cytostatic;  
 KM antiproliferative; anticancer; antipsoriatic; ss.  
 XX  
 OS Unidentified.  
 XX  
 PN US6068982-A.  
 XX  
 PD 30-MAY-2000.  
 XX  
 PF 17-DEC-1996; 96US-00767942.  
 XX  
 PR 04-JAN-1994; 94US-00176937.  
 PR 23-MAY-1994; 94US-00247904.  
 PR 27-MAY-1994; 94US-00250795.  
 PR 13-SEP-1994; 94US-00305520.  
 PR 07-JUN-1995; 95US-00486663.  
 XX  
 PA (MITO-) MITOTIX INC.  
 XX  
 PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damaguer V, Rolfe M;  
 XX  
 DR WPI; 2000-410854/35.  
 XX  
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory  
 PT protein for treating cancers involves measuring ubiquitination levels of  
 PT the protein in the presence of candidate agent in an eukaryotic cell.  
 XX  
 PS Example 7; Fig 5; 73pp; English.  
 XX  
 CC The invention relates to a method of identifying an inhibitor of  
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein  
 CC comprising contacting an engineered eukaryotic cell with a candidate  
 CC agent. The eukaryotic cells is engineered to express a recombinant human,  
 CC Candida albicans or Schizosaccharomyces pombe ubiquitin-conjugating  
 CC enzyme (AAB03169-B03171), a cell cycle regulatory protein (such as p53)  
 CC and ubiquitin. The specification also discloses novel Candida albicans  
 CC and Schizosaccharomyces pombe ubiquitin-conjugating enzymes, CAUBC and  
 CC SPUBCR (AAB03170, AAB03171), and two novel human ubiquitin-conjugating  
 CC enzymes, hUBC and rapUBC (AAB03169, AAB03173). The ubiquitin-mediated  
 CC proteolysis system is the major pathway for the selective, controlled  
 CC degradation of intracellular proteins in eukaryotic cells. In particular,  
 CC this system controls the half-lives of cellular proteins, and is  
 CC important in controlling the levels of proteins involved in cell cycle  
 CC progression. Alterations in the ubiquitination of these proteins may  
 CC therefore play a role in the development of cancers. For example, human  
 CC papillomaviruses such as HPV-18 encode a transforming protein, E6  
 CC (AAB03176), which combines with a cellular B6-associated protein (B6-AP;  
 CC AAB03177) to stimulate the ubiquitination of p53, thus targeting it for  
 CC degradation. The ubiquitination inhibitors identified according to the  
 CC method of the invention are useful for treatment of cervical cancers and  
 CC connective tissue disorders and for controlling the wound healing  
 CC process. They are also useful in treatment of hyperplastic epidermal  
 CC conditions such as psoriasis, neoplastic epidermal conditions, skin  
 CC cancers e.g., basal cell carcinoma, squamous cell carcinoma. The  
 CC inhibitors are useful for deregulating myc expression and rendering the  
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of  
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-  
 CC mediated degradation of cyclins are useful as antiproliferative agents.  
 CC Sequences AAB61639-A61640 represent opposite strands of a p53 binding  
 CC site oligonucleotide which may be incorporated in a reporter construct  
 CC for use in an assay of the invention  
 XX  
 SQ Sequence 26 BP; 4 A; 10 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGATGTC 20  
 |||||  
 DB 6 GGACATGCCCGGCGATGTC 25

RESULT 20  
 AAA61639/c  
 ID AAA61639 standard; DNA; 26 BP.  
 XX  
 AC AAA61639;  
 XX  
 DT 23-OCT-2000 (first entry)  
 XX  
 DE p53 binding site oligonucleotide, SEQ ID NO:40.  
 XX  
 KM p53 binding site; ubiquitin conjugating enzyme;  
 KM ubiquitin mediated proteolysis; human; cellular protein half life;  
 KM ubiquitination inhibitor; p53; cyclin; cell cycle regulator;  
 KM myc deregulation; human papillomavirus; HPV-18 B6 protein;  
 KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;  
 KM psoriasis; connective tissue disorder; wound healing; cytostatic;  
 KM antiproliferative; anticancer; antipsoriatic; ss.  
 XX  
 OS Unidentified.  
 XX  
 PN US6068982-A.  
 XX  
 PD 30-MAY-2000.  
 XX  
 PF 17-DEC-1996; 96US-00767942.  
 XX  
 PR 04-JAN-1994; 94US-00176937.  
 PR 23-MAY-1994; 94US-00247904.  
 PR 27-MAY-1994; 94US-00250795.  
 PR 13-SEP-1994; 94US-00305520.  
 PR 07-JUN-1995; 95US-00486663.  
 XX  
 PA (MITO-) MITOTIX INC.  
 XX  
 PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M;  
 XX  
 DR WPI; 2000-410854/35.  
 XX  
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory  
 PT protein for treating cancers involves measuring ubiquitination levels of  
 PT the protein in the presence of candidate agent in an eukaryotic cell.  
 XX  
 PS Example 7, Fig 5, 73pp; English.  
 XX  
 CC The invention relates to a method of identifying an inhibitor of  
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein  
 CC comprising contacting an engineered eukaryotic cell with a candidate  
 CC agent. The eukaryotic cells is engineered to express a recombinant human,  
 CC Candida albicans or Schizosaccharomyces pombe ubiquitin- conjugating  
 CC enzyme (AAB03169-B03171), a cell cycle regulatory protein (such as p53)  
 CC and ubiquitin. The specification also discloses novel Candida albicans  
 CC and Schizosaccharomyces pombe ubiquitin-conjugating enzymes, cab0cs and  
 CC sub0ce (AAB03170, AAB03171), and two novel human ubiquitin-conjugating  
 CC enzymes, hUBC8 and rapBC (AAB03169, AAB03173). The ubiquitin-mediated  
 CC proteolysis system is the major pathway for the selective, controlled  
 CC degradation of intracellular proteins in eukaryotic cells. In particular,  
 CC this system controls the half-lives of cellular proteins, and is  
 CC important in controlling the levels of proteins involved in cell cycle  
 CC progression. Alterations in the ubiquitination of these proteins may  
 CC therefore play a role in the development of cancers. For example, human  
 CC papillomaviruses such as HPV-18 encode a transforming protein, B6  
 CC (AAB03176), which combines with a cellular B6-associated protein (B6-AP;  
 CC AAB03177) to stimulate the ubiquitination of p53, thus targeting it for  
 CC degradation. The ubiquitination inhibitors identified according to the

CC method of the invention are useful for treatment of cervical cancers and  
 CC connective tissue disorders and for controlling the wound healing  
 CC process. They are also useful in treatment of hyperplastic epidermal  
 CC conditions such as psoriasis, neoplastic epidermal conditions, skin  
 CC cancers e.g., basal cell carcinomas, squamous cell carcinomas. The  
 CC inhibitors are useful for deregulating myc expression and rendering the  
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of  
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-  
 CC mediated degradation of cyclins are useful as antiproliferative agents.  
 CC Sequences AAA61639-A61640 represent opposite strands of a p53 binding  
 CC site oligonucleotide which may be incorporated in a reporter construct  
 CC for use in an assay of the invention  
 XX  
 SQ Sequence 26 BP; 4 A; 10 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGATGTC 20  
 |||||  
 DB 25 GGACATGCCCGGCGATGTC 6

RESULT 21  
 AAA61640  
 ID AAA61640 standard; DNA; 27 BP.  
 XX  
 AC AAA61640;  
 XX  
 DT 23-OCT-2000 (first entry)  
 XX  
 DE p53 binding site oligonucleotide, SEQ ID NO:41.  
 XX  
 KM p53 binding site; ubiquitin conjugating enzyme;  
 KM ubiquitin mediated proteolysis; human; cellular protein half life;  
 KM ubiquitination inhibitor; p53; cyclin; cell cycle regulator;  
 KM myc deregulation; human papillomavirus; HPV-18 B6 protein;  
 KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;  
 KM psoriasis; connective tissue disorder; wound healing; cytostatic;  
 KM antiproliferative; anticancer; antipsoriatic; ss.  
 XX  
 OS Unidentified.  
 XX  
 PN US6068982-A.  
 XX  
 PD 30-MAY-2000.  
 XX  
 PF 17-DEC-1996; 96US-00767942.  
 XX  
 PR 04-JAN-1994; 94US-00176937.  
 PR 23-MAY-1994; 94US-00247904.  
 PR 27-MAY-1994; 94US-00250795.  
 PR 13-SEP-1994; 94US-00305520.  
 PR 07-JUN-1995; 95US-00486663.  
 XX  
 PA (MITO-) MITOTIX INC.  
 XX  
 PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M;  
 XX  
 DR WPI; 2000-410854/35.  
 XX  
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory  
 PT protein for treating cancers involves measuring ubiquitination levels of  
 PT the protein in the presence of candidate agent in an eukaryotic cell.  
 XX  
 PS Example 7, Fig 5, 73pp; English.  
 XX  
 CC The invention relates to a method of identifying an inhibitor of  
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein  
 CC comprising contacting an engineered eukaryotic cell with a candidate  
 CC agent. The eukaryotic cells is engineered to express a recombinant human,  
 CC Candida albicans or Schizosaccharomyces pombe ubiquitin- conjugating

CC enzyme (AAB03169-B03171), a cell cycle regulatory protein (such as p53)  
 CC and ubiquitin. The specification also discloses novel *Candida albicans*  
 CC and *Schizosaccharomyces pombe* ubiquitin-conjugating enzymes, caubCC and  
 CC spubCC (AAB03170, AAB03171), and two novel human ubiquitin-conjugating  
 CC enzymes, hHubCC and rapubCC (AAB03169, AAB03173). The ubiquitin-mediated  
 CC proteolysis system is the major pathway for the selective, controlled  
 CC degradation of intracellular proteins in eukaryotic cells. In particular,  
 CC this system controls the half-lives of cellular proteins, and is  
 CC important in controlling the levels of proteins involved in cell cycle  
 CC progression. Alterations in the ubiquitination of these proteins may  
 CC therefore play a role in the development of cancers. For example, human  
 CC papillomaviruses such as HPV-18 encode a transforming protein, E6  
 CC (AAB03176), which combines with a cellular E6-associated protein (E6-AP;  
 CC AAB03177) to stimulate the ubiquitination of p53, thus targeting it for  
 CC degradation. The ubiquitination inhibitors identified according to the  
 CC method of the invention are useful for treatment of cervical cancers and  
 CC connective tissue disorders and for controlling the wound healing  
 CC process. They are also useful in treatment of hyperplastic epidermal  
 CC conditions such as psoriasis, neoplastic epidermal conditions, skin  
 CC cancers e.g., basal cell carcinomas, squamous cell carcinomas. The  
 CC inhibitors are useful for deregulating myc expression and rendering the  
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of  
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-  
 CC mediated degradation of cyclins are useful as antiproliferative agents.  
 CC Sequences AAB61639-A61640 represent opposite strands of a p53 binding  
 CC site oligonucleotide which may be incorporated in a reporter construct  
 CC for use in an assay of the invention  
 CC XX

Sequence 27 BP; 4 A; 9 C; 10 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 Db 7 GGACATGCCCGGCATGTCC 26

RESULT 22

AAA61640/c  
 ID AAA61640 standard; DNA; 27 BP.

XX AAA61640;

DT 23-OCT-2000 (first entry)

DE p53 binding site oligonucleotide, SEQ ID NO:41.

XX p53 binding site; ubiquitin conjugating enzyme;

KM ubiquitin mediated proteolysis; human; cellular protein half life;

KM ubiquitination inhibitor; p53; cyclin; cell cycle regulator;

KM myc deregulation; human papillomavirus; HPV-18 E6 protein;

KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;

KM psoriasis; connective tissue disorder; wound healing; cytostatic;

KM antiproliferative; anticancer; antiproliferative; ss.

XX Unidentified.

OS US6068982-A.

PN 30-MAY-2000.

PF 17-DEC-1996; 96US-00767942.

XX 04-JAN-1994; 94US-00176937.

PR 23-MAY-1994; 94US-00247904.

PR 27-MAY-1994; 94US-00250795.

PR 13-SEP-1994; 94US-00305520.

PR 07-JUN-1995; 95US-00486663.

XX (MITO-) MITOTIX INC.

XX PA

PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damagmez V, Rolfe M;  
 XX WPI; 2000-410854/35.  
 DR  
 XX  
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory  
 PT protein for treating cancers involves measuring ubiquitination levels of  
 PT the protein in the presence of candidate agent in an eukaryotic cell.  
 XX  
 PS Example 7, Fig 5; 73pp; English.

XX The invention relates to a method of identifying an inhibitor of  
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein  
 CC comprising contacting an engineered eukaryotic cell with a candidate  
 CC agent. The eukaryotic cells is engineered to express a recombinant human,  
 CC *Candida albicans* or *Schizosaccharomyces pombe* ubiquitin-conjugating  
 CC enzyme (AAB03169-B03171), a cell cycle regulatory protein (such as p53)  
 CC and ubiquitin. The specification also discloses novel *Candida albicans*  
 CC and *Schizosaccharomyces pombe* ubiquitin-conjugating enzymes, caubCC and  
 CC spubCC (AAB03170, AAB03171), and two novel human ubiquitin-conjugating  
 CC enzymes, hHubCC and rapubCC (AAB03169, AAB03173). The ubiquitin-mediated  
 CC proteolysis system is the major pathway for the selective, controlled  
 CC degradation of intracellular proteins in eukaryotic cells. In particular,  
 CC this system controls the half-lives of cellular proteins, and is  
 CC important in controlling the levels of proteins involved in cell cycle  
 CC progression. Alterations in the ubiquitination of these proteins may  
 CC therefore play a role in the development of cancers. For example, human  
 CC papillomaviruses such as HPV-18 encode a transforming protein, E6  
 CC (AAB03176), which combines with a cellular E6-associated protein (E6-AP;  
 CC AAB03177) to stimulate the ubiquitination of p53, thus targeting it for  
 CC degradation. The ubiquitination inhibitors identified according to the  
 CC method of the invention are useful for treatment of cervical cancers and  
 CC connective tissue disorders and for controlling the wound healing  
 CC process. They are also useful in treatment of hyperplastic epidermal  
 CC conditions such as psoriasis, neoplastic epidermal conditions, skin  
 CC cancers e.g., basal cell carcinomas, squamous cell carcinomas. The  
 CC inhibitors are useful for deregulating myc expression and rendering the  
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of  
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-  
 CC mediated degradation of cyclins are useful as antiproliferative agents.  
 CC Sequences AAB61639-A61640 represent opposite strands of a p53 binding  
 CC site oligonucleotide which may be incorporated in a reporter construct  
 CC for use in an assay of the invention  
 CC XX

Sequence 27 BP; 4 A; 9 C; 10 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
 |||||  
 Db 26 GGACATGCCCGGCATGTCC 7

RESULT 23

AAA38055 standard; DNA; 56 BP.

XX AAA38055;

DT 24-AUG-2000 (first entry)

DE linker used in p53CON-luciferase plasmid construction.

XX Selective replication; recombinant virus; pathway responsive promoter;

KM viral replication repressor; cell death; neoplastic ablate; TGF-beta;

KM transforming growth factor-beta; anti-proliferative; tumour; cancer;

XX p53CON; luciferase; ds.

XX Synthetic.

OS WO200022137-A2.

XX

XX

```

PD 20-APR-2000.
PF 14-OCT-1999; 99MO-US021452.
PR 15-OCT-1999; 98US-00172686.
PA (CANU-) CANU1 INC.
XX Ramachandra M, Shabram PW;
XX MPI; 2000-317990/27.
DR
XX
XX
XX Recombinant viruses which selectively replicate in target cells used in
PT the treatment of tumors comprise a pathway-responsive promoter linked to
PT a viral replication repressor.
XX
XX
PS Example 3; Page 37; 49pp; English.
XX
XX The present invention relates to a selectively replicating recombinant
CC virus, comprising a pathway responsive promoter, linked to a repressor of
CC viral replication. This sequence represents a linker used in the
CC construction of a luciferase plasmid with p53-responsive promoter p53CON.
CC Replication of the recombinant virus responds to the intracellular
CC conditions of the target cell through the use of a pathway-responsive
CC promoter driving expression of an inhibitor of viral replication. In the
CC target cell, the promoter element is inactive and viral replication is
CC permitted. The recombinant virus is used for killing a target cell in
CC vivo with pathway defect on contact or ex vivo to eliminate tumour cells
CC from stem cell products. The virus is also used for ablating neoplastic
CC cells in a mammalian organism preferably humans, pigs, horses, cattle,
CC dogs or cats in vivo. A recombinant adenoviral vector containing multiple
CC promoters such that the replication of the virus is selectively enabled
CC in cells possessing multiple pathway defects is useful in the treatment
CC of tumours associated with a lack of TGF-beta antiproliferative action.
CC The tumours include breast carcinomas, hepatomas, gastric, colon and skin
CC tumours, as well as B and T lymphomas. The vectors can also be used for
CC the treatment of tumour cells containing p53-pathway defects. They are
CC also useful in treating cancer cells with Rb pathway defects.
CC Alternatively, such vectors may be employed to express a cell surface
CC protein capable of recognition by a binding molecule such as a
CC fluorescently labelled antibody
XX
SQ Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 3; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGTCC 20
Db 7 GGACATGCCCGGCGCATGTCC 26
RESULT 24
ID AAA38055/c
XX AAA38055 standard; DNA; 56 BP.
XX
XX AAA38055;
XX
XX 24-AUG-2000 (first entry)
XX
XX Linker used in p53CON-luciferase plasmid construction.
DB Selective replication; recombinant virus; pathway responsive promoter;
XX viral replication repressor; cell death; neoplastic; ablute; TGF-beta;
XX transforming growth factor-beta; anti-proliferative; tumour; cancer;
XX p53CON; luciferase; ds.
XX
XX Synthetic.
XX
XX WO2000022137-A2.
XX
XX 20-APR-2000.

```

XX  
PF 14-OCT-1999; 99WO-US021452.  
XX  
PR 15-OCT-1998; 98US-00172686.  
XX  
PA (CANU-) CANUI INC.  
XX Ramachandra M, Shabram PW;  
XX MPI; 2000-317990/27.  
XX  
PT Recombinant viruses which selectively replicate in target cells used in  
PT the treatment of tumors comprise a pathway-responsive promoter linked to  
PT a viral replication repressor.  
XX  
PS Example 3; Page 37; 49pp; English.  
XX  
CC The present invention relates to a selectively replicating recombinant  
CC virus, comprising a pathway responsive promoter, linked to a repressor of  
CC viral replication. This sequence represents a linker used in the  
CC construction of a luciferase plasmid with p53-responsive promoter p53CON.  
CC Replication of the recombinant virus responds to the intracellular  
CC conditions of the target cell through the use of a pathway-responsive  
CC promoter driving expression of an inhibitor of viral replication. In the  
CC target cell, the promoter element is inactive and viral replication is  
CC permitted. The recombinant virus is used for killing a target cell in  
CC vivo with pathway defect on contact or ex vivo to eliminate tumour cells  
CC from stem cell products. The virus is also used for ablating neoplastic  
CC cells in a mammalian organism preferably humans, pigs, horses, cattle,  
CC dogs or cats in vivo. A recombinant adenoviral vector containing multiple  
CC promoters such that the replication of the virus is selectively enabled  
CC in cells possessing multiple pathway defects is useful in the treatment  
CC of tumours associated with a lack of TGF-beta antiproliferative action.  
CC The tumours include breast carcinomas, hepatomas, gastric, colon and skin  
CC tumours, as well as B and T lymphomas. The vectors can also be used for  
CC the treatment of tumour cells containing p53-pathway defects. They are  
CC also useful in treating cancer cells with Kd pathway defects.  
CC Alternatively, such vectors may be employed to express a cell surface  
CC protein capable of recognition by a binding molecule such as a  
CC fluorescently labelled antibody

SQ Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 56;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
|||||  
DB 51 GGACATGCCCGGCATGTCC 32

RESULT 25  
ABX16465  
ID ABX16465 standard; DNA; 56 BP.

XX AC ABX16465;  
XX  
DT 15-APR-2003 (first entry)  
XX DE Selectively replicating virus 01/PEME associated PCR primer #5.  
XX  
KW Selectively replicating virus; pathway-responsive promoter;  
KW repressor of viral replication; p53-pathway defect;  
KW transforming growth factor beta; TGFbeta; antiproliferative;  
KW breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;  
KW PCR; primer; ss; 01/PEME.  
XX  
OS Synthetic.  
XX  
PN US2002150557-A1.  
XX  
PD 17-OCT-2002.

```

XX 30-JAN-2002; 2002US-00062216.
PF 15-OCT-1998; 98US-0104399P.
PR 13-OCT-1999; 99US-00416812.
XX (RAMA/) RAMACHANDRA M.
PA (SHAB/) SHABRAM P W.
XX Ramachandra M, Shabram PW;
XX WPI, 2003-198263/19.
DR New recombinant virus utilizing a pathway-responsive promoter that
PT inhibits viral replication, useful for diagnosing and treating cancer
PT with p53-pathway defects or with a lack of TGFBproximatelyb
PT antiproliferative action.
XX Example 11; Page 23; 51pp; English.
XX The invention describes a selective replicating recombinant virus (I)
CC comprising a pathway-responsive promoter operably linked to a repressor
CC of viral replication. The methods and compositions are useful for the
CC diagnosis and treatment of disorders containing p53-pathway defects or
CC disorders associated with a lack of transforming growth factor beta
CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
CC and skin cancer. This sequence represents a PCR primer used in the
CC creation of selectively replicating virus 01/PEMB
XX Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 8; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGACATGCCCGGCATGTCC 20
Db 7 GGACATGCCCGGCATGTCC 26
RESULT 26
ABX16465/C
ID ABX16465 standard; DNA; 56 BP.
XX ABX16465;
AC 15-APR-2003 (first entry)
XX 15-APR-2003 (first entry)
XX Selectively replicating virus 01/PEMB associated PCR primer #5.
DE Selectively replicating virus; pathway-responsive promoter;
XX repressor of viral replication; p53-pathway defect;
XX transforming growth factor beta; TGFbeta; antiproliferative;
XX breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX PCR; primer; ss; 01/PEMB.
XX Synthetic.
XX US2002150557-A1.
XX 17-OCT-2002.
XX 30-JAN-2002; 2002US-00062216.
PF 15-OCT-1998; 98US-0104399P.
PR 13-OCT-1999; 99US-00416812.
XX (RAMA/) RAMACHANDRA M.
PA (SHAB/) SHABRAM P W.
XX Ramachandra M, Shabram PW;
XX WPI, 2003-198263/19.

```

```

XX New recombinant virus utilizing a pathway-responsive promoter that
PT inhibits viral replication, useful for diagnosing and treating cancer
PT with p53-pathway defects or with a lack of TGFBproximatelyb
PT antiproliferative action.
XX Example 11; Page 23; 51pp; English.
XX The invention describes a selective replicating recombinant virus (I)
CC comprising a pathway-responsive promoter operably linked to a repressor
CC of viral replication. The methods and compositions are useful for the
CC diagnosis and treatment of disorders containing p53-pathway defects or
CC disorders associated with a lack of transforming growth factor beta
CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
CC and skin cancer. This sequence represents a PCR primer used in the
CC creation of selectively replicating virus 01/PEMB
XX Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 8; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGACATGCCCGGCATGTCC 20
Db 51 GGACATGCCCGGCATGTCC 32
RESULT 27
ABX16450
ID ABX16450 standard; DNA; 56 BP.
XX ABX16450;
AC 15-APR-2003 (first entry)
XX 15-APR-2003 (first entry)
XX p53CON-luciferase plasmid associated PCR primer #2.
DE Selectively replicating virus; pathway-responsive promoter;
XX repressor of viral replication; p53-pathway defect;
XX transforming growth factor beta; TGFbeta; antiproliferative;
XX breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.
XX Synthetic.
XX US2002150557-A1.
XX 17-OCT-2002.
XX 30-JAN-2002; 2002US-00062216.
PF 15-OCT-1998; 98US-0104399P.
PR 13-OCT-1999; 99US-00416812.
XX (RAMA/) RAMACHANDRA M.
PA (SHAB/) SHABRAM P W.
XX Ramachandra M, Shabram PW;
XX WPI, 2003-198263/19.
XX New recombinant virus utilizing a pathway-responsive promoter that
PT inhibits viral replication, useful for diagnosing and treating cancer
PT with p53-pathway defects or with a lack of TGFBproximatelyb
PT antiproliferative action.
XX Example 3; Page 21; 51pp; English.
XX The invention describes a selective replicating recombinant virus (I)
CC comprising a pathway-responsive promoter operably linked to a repressor
CC of viral replication. The methods and compositions are useful for the
CC diagnosis and treatment of disorders containing p53-pathway defects or

```



CC disorders associated with a lack of transforming growth factor beta  
CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon  
CC and skin cancer. This sequence represents a PCR primer used in the  
CC creation of the p53CON-luciferase plasmid containing a p53-responsive  
CC responsive promoter

CC Sequence 56 BP; 10 A; 17 C; 20 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 8; Length 56;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20  
Db 2 GGACATGCCCGGCATGTCC 21

RESULT 28  
ABX16450/c  
ID ABX16450 standard; DNA; 56 BP.

XX AC ABX16450;  
XX 15-APR-2003 (first entry)

XX DE p53CON-luciferase plasmid associated PCR primer #2.

XX KM Selectively replicating virus; pathway-responsive promoter;  
XX KM repressor of viral replication; p53-pathway defect;  
XX KM transforming growth factor beta; TGFbeta; antiproliferative;  
XX KM breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;  
XX KM p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.

XX OS Synthetic.

XX PN US2002150557-A1.

XX PD 17-OCT-2002.

XX PF 30-JAN-2002; 2002US-00062216.

XX PR 15-OCT-1998; 98US-0104399P.

XX PR 13-OCT-1999; 99US-00416812.

XX PA (RAMA/) RAMACHANDRA M.

XX PA (SHAB/) SHABRAM P W.

XX PI Ramachandra M, Shabram PW;

XX DR WPI; 2003-198263/19.

XX PT New recombinant virus utilizing a pathway-responsive promoter that  
XX PT inhibits viral replication, useful for diagnosing and treating cancer  
XX PT with p53-pathway defects or with a lack of TGFaproximately  
XX PT antiproliferative action.

XX PS Example 3; Page 21; 51pp; English.

XX CC The invention describes a selective replicating recombinant virus (I)  
XX CC comprising a pathway-responsive promoter operably linked to a repressor  
XX CC of viral replication. The methods and compositions are useful for the  
XX CC diagnosis and treatment of disorders containing p53-pathway defects or  
XX CC disorders associated with a lack of transforming growth factor beta  
XX CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon  
XX CC and skin cancer. This sequence represents a PCR primer used in the  
XX CC creation of the p53CON-luciferase plasmid containing a p53-responsive  
XX CC responsive promoter

XX SQ Sequence 56 BP; 10 A; 17 C; 20 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 8; Length 56;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20  
Db 46 GGACATGCCCGGCATGTCC 27

RESULT 29  
ABX16449  
ID ABX16449 standard; DNA; 56 BP.

XX AC ABX16449;

XX DT 15-APR-2003 (first entry)

XX DE p53CON-luciferase plasmid associated PCR primer #1.

XX KM Selectively replicating virus; pathway-responsive promoter;  
XX KM repressor of viral replication; p53-pathway defect;  
XX KM transforming growth factor beta; TGFbeta; antiproliferative;  
XX KM breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;  
XX KM p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.

XX OS Synthetic.

XX PN US2002150557-A1.

XX PD 17-OCT-2002.

XX PF 30-JAN-2002; 2002US-00062216.

XX PR 15-OCT-1998; 98US-0104399P.

XX PR 13-OCT-1999; 99US-00416812.

XX PA (RAMA/) RAMACHANDRA M.

XX PA (SHAB/) SHABRAM P W.

XX PI Ramachandra M, Shabram PW;

XX DR WPI; 2003-198263/19.

XX PT New recombinant virus utilizing a pathway-responsive promoter that  
XX PT inhibits viral replication, useful for diagnosing and treating cancer  
XX PT with p53-pathway defects or with a lack of TGFaproximately  
XX PT antiproliferative action.

XX PS Example 3; Page 21; 51pp; English.

XX CC The invention describes a selective replicating recombinant virus (I)  
XX CC comprising a pathway-responsive promoter operably linked to a repressor  
XX CC of viral replication. The methods and compositions are useful for the  
XX CC diagnosis and treatment of disorders containing p53-pathway defects or  
XX CC disorders associated with a lack of transforming growth factor beta  
XX CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon  
XX CC and skin cancer. This sequence represents a PCR primer used in the  
XX CC creation of the p53CON-luciferase plasmid containing a p53-responsive  
XX CC responsive promoter

XX SQ Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 8; Length 56;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20  
Db 7 GGACATGCCCGGCATGTCC 26

RESULT 30  
ABX16449/c  
ID ABX16449 standard; DNA; 56 BP.

XX AC ABX16449;

XX	15-APR-2003	(first entry)
DT		
XX		
DE	p53CON-luciferase plasmid associated PCR primer #1.	
XX		
KM	selectively replicating virus; pathway-responsive promoter;	
KM	repressor of viral replication; p53-pathway defect;	
KM	transforming growth factor beta; TGFbeta; antiproliferative;	
KM	breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;	
KM	p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.	
OS	Synthetic.	
XX		
PN	US2002150557-A1.	
XX		
PD	17-OCT-2002.	
XX		
PF	30-JAN-2002; 2002US-00062216.	
XX		
PR	15-OCT-1998; 98US-0104399P.	
XX		
PR	13-OCT-1999; 99US-00416812.	
XX		
PA	(RAMA/) RAMACHANDRA M.	
XX	(SHAB/) SHABRAM P W.	
PI		
XX	Ramachandra M, Shabram PW;	
DR	WPI, 2003-198263/19.	
XX		
PT	New recombinant virus utilizing a pathway-responsive promoter that	
XX	inhibits viral replication, useful for diagnosing and treating cancer	
PT	with p53-pathway defects or with a lack of TGFapproximately	
XX	antiproliferative action.	
PS		
XX	Example 3; Page 21; 51pp; English.	
XX		
CC	The invention describes a selective replicating recombinant virus (I)	
CC	comprising a pathway-responsive promoter operably linked to a repressor	
CC	of viral replication. The methods and compositions are useful for the	
CC	diagnosis and treatment of disorders containing p53-pathway defects or	
CC	(Tgfbeta) associated with a lack of transforming growth factor beta	
CC	(Tgfbeta) antiproliferative action, such as breast, liver, gastric, colon	
CC	and skin cancer. This sequence represents a PCR primer used in the	
CC	creation of the p53CON-luciferase plasmid containing a p53-responsive	
CC	responsive promoter	
XX		
SQ	Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;	
XX		
Query Match	100.0%; Score 20; DB 8; Length 56;	
Beet Local Similarity	100.0%; Pred. No. 6.7;	
Matches 20; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GGACATGCCCGGGCATGTCC 20	
DB		
	51 GGACATGCCCGGGCATGTCC 32	
XX		
RESULT 31		
ABX16466		
ID	ABX16466 standard; DNA; 56 BP.	
XX		
AC	ABX16466;	
XX		
DT	15-APR-2003 (first entry)	
XX		
DE	selectively replicating virus 01/PEME associated PCR primer #6.	
XX		
KM	selectively replicating virus; pathway-responsive promoter;	
KM	repressor of viral replication; p53-pathway defect;	
KM	transforming growth factor beta; TGFbeta; antiproliferative;	
KM	breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;	
KM	breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;	
KM	PCR; primer; ss; 01/PEME.	
XX		

OS	Synthetic.
XX	
PN	US2002150557-A1.
XX	
PD	17-OCT-2002.
XX	
PF	30-JAN-2002; 2002US-00062216.
XX	
PR	15-OCT-1998; 98US-0104399P.
PR	13-OCT-1999; 99US-00416812.
XX	
PA	(RAMA/) RAMACHANDRA M.
PA	(SHAB/) SHABRAM P W.
XX	
PI	Ramachandra M, Shabram PW;
XX	
DR	WPI; 2003-198263/19.
XX	
PT	New recombinant virus utilizing a pathway-responsive promoter that
PT	inhibits viral replication, useful for diagnosing and treating cancer
PT	with p53-pathway defects or with a lack of TGFapproximately
PT	antiproliferative action.
XX	
PS	Example 11; Page 23-24; 51pp; English.
XX	
CC	The invention describes a selective replicating recombinant virus (I)
CC	comprising a pathway-responsive promoter operably linked to a repressor
CC	of viral replication. The methods and compositions are useful for the
CC	diagnosis and treatment of disorders containing p53-pathway defects or
CC	disorders associated with a lack of transforming growth factor beta
CC	(TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
CC	and skin cancer. This sequence represents a PCR primer used in the
CC	creation of selectively replicating virus 01/PEME
XX	
SQ	Sequence 56 BP; 10 A; 17 C; 19 G; 9 T; 0 U; 1 Other;
QY	
Query Match	100.0%; Score 20; DB 8; Length 56;
Best Local Similarity	100.0%; Pred. NO. 6.7;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	
1	GGACATGCCCGGCGCATGTC 20
2	GGACATGCCCGGCGCATGTC 21
RESULT 32	
ABX16466/C	
ID	ABX16466 standard; DNA; 56 BP.
XX	
AC	ABX16466;
XX	
DT	15-APR-2003 (first entry)
XX	
DE	Selectively replicating virus 01/PEME associated PCR primer #6.
XX	
KM	Selectively replicating virus; pathway-responsive promoter;
KM	repressor of viral replication; p53-pathway defect;
KM	transforming growth factor beta; TGFbeta; antiproliferative;
KM	breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
KM	PCR; primer; ss; 01/PEME.
XX	
OS	Synthetic.
XX	
PN	US2002150557-A1.
XX	
PD	17-OCT-2002.
XX	
PF	30-JAN-2002; 2002US-00062216.
XX	
PR	15-OCT-1998; 98US-0104399P.
PR	13-OCT-1999; 99US-00416812.
XX	
PA	(RAMA/) RAMACHANDRA M.

PA (SHAB/) SHABRAM P W.  
 XX Ramachandra M, Shabram PW;  
 XX WPI; 2003-198263/19.  
 XX  
 PT New recombinant virus utilizing a pathway-responsive promoter that  
 PT inhibits viral replication, useful for diagnosing and treating cancer  
 PT with p53-pathway defects or with a lack of TGFBpropproximately  
 PT antiproliferative action.  
 XX  
 XX Example 11; Page 23-24; 51pp; English.  
 XX  
 CC The invention describes a selective replicating recombinant virus (I)  
 CC comprising a pathway-responsive promoter operably linked to a repressor  
 CC of viral replication. The methods and compositions are useful for the  
 CC diagnosis and treatment of disorders containing p53-pathway defects or  
 CC disorders associated with a lack of transforming growth factor beta  
 CC (TGFBeta) antiproliferative action, such as breast, liver, gastric, colon  
 CC and skin cancer. This sequence represents a PCR primer used in the  
 CC creation of selectively replicating virus 01/PPME  
 XX  
 SQ Sequence 56 BP; 10 A; 17 C; 19 G; 9 T; 0 U; 1 Other;  
 Query Match 100.0%; Score 20; DB 8; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 6.7;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCGCATGTCC 20  
 Db 46 GGACATGCCCGGCGCATGTCC 27  
 RESULT 33  
 ADE52736  
 ID ADE52736 standard; DNA; 19 BP.  
 XX  
 AC ADE52736;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Oligonucleotide SEQ ID 102.  
 XX  
 KM DNA-binding protein; interferon-activatable protein; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2003089466-A1.  
 XX  
 PD 30-OCT-2003.  
 XX  
 PF 18-APR-2003; 2003WO-JP004981.  
 XX  
 PR 19-APR-2002; 2002JP-00117840.  
 PR 30-APR-2002; 2002JP-00128418.  
 PR 30-APR-2002; 2002JP-00128779.  
 PR 04-DEC-2002; 2002JP-00352469.  
 XX  
 PA (RIKE) RIKEN KK.  
 PA (DNAF-) DNAFORM KK.  
 PA (MITU) MITSUBISHI CHEM CORP.  
 XX  
 PI Hayashizaki Y, Kamiya M, Kubodera H;  
 XX WPI; 2004-011681/01.  
 XX  
 DR WPI; 2004-011681/01.  
 XX  
 PT Proteins with DNA binding activity and substances that affect their  
 PT activity or expression, useful for treating associated disorders.  
 XX  
 PS Example 9; SEQ ID NO 102; 237pp; Japanese.  
 XX  
 CC The present invention relates to novel proteins (ADE52648-ADE52660,  
 CC ADE52670 and ADE52672) and their coding sequences (ADE52635-ADE52647,  
 CC ADE52670 and ADE52672) and their coding sequences (ADE52635-ADE52647,

CC ADE52669 and ADE52671). The proteins have a DNA-binding activity or an  
 CC interferon-activatable protein (IAP)-like activity. The present  
 CC oligonucleotide is related to p53.  
 XX  
 SQ Sequence 19 BP; 3 A; 6 C; 7 G; 3 T; 0 U; 0 Other;  
 Query Match 95.0%; Score 19; DB 12; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCGCATGTTC 19  
 Db 1 GGACATGCCCGGCGCATGTTC 19  
 RESULT 34  
 ADE52736/C  
 ID ADE52736 standard; DNA; 19 BP.  
 XX  
 AC ADE52736;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Oligonucleotide SEQ ID 102.  
 XX  
 KM DNA-binding protein; interferon-activatable protein; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2003089466-A1.  
 XX  
 PD 30-OCT-2003.  
 XX  
 PF 18-APR-2003; 2003WO-JP004981.  
 XX  
 PR 19-APR-2002; 2002JP-00117840.  
 PR 30-APR-2002; 2002JP-00128418.  
 PR 30-APR-2002; 2002JP-00128779.  
 PR 04-DEC-2002; 2002JP-00352469.  
 XX  
 PA (RIKE) RIKEN KK.  
 PA (DNAF-) DNAFORM KK.  
 PA (MITU) MITSUBISHI CHEM CORP.  
 XX  
 PI Hayashizaki Y, Kamiya M, Kubodera H;  
 XX WPI; 2004-011681/01.  
 XX  
 DR WPI; 2004-011681/01.  
 XX  
 PT Proteins with DNA binding activity and substances that affect their  
 PT activity or expression, useful for treating associated disorders.  
 XX  
 PS Example 9; SEQ ID NO 102; 237pp; Japanese.  
 XX  
 CC The present invention relates to novel proteins (ADE52648-ADE52660,  
 CC ADE52670 and ADE52672) and their coding sequences (ADE52635-ADE52647,  
 CC ADE52669 and ADE52671). The proteins have a DNA-binding activity or an  
 CC interferon-activatable protein (IAP)-like activity. The present  
 CC oligonucleotide is related to p53.  
 XX  
 SQ Sequence 19 BP; 3 A; 6 C; 7 G; 3 T; 0 U; 0 Other;  
 Query Match 95.0%; Score 19; DB 12; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GGACATGCCCGGCGCATGTTC 20  
 Db 19 GGACATGCCCGGCGCATGTTC 1  
 RESULT 35  
 AAT32849  
 ID AAT32849 standard; DNA; 30 BP.

```

XX AC AAT32849;
XX XX
XX 06-NOV-1996 (first entry)
XX DT
XX XX
XX DE Tumour suppressor p53 binding site oligonucleotide BC.
XX XX
XX p53 protein; tumour suppressor; tetramerisation domain;
XX KW chimaeric protein; GCN4; transcription factor; leucine zipper;
XX KW gene therapy; vector; cell proliferation; cancer; apoptosis;
XX KW autoimmune disease; immune tolerance; ds.
XX XX
XX OS Synthetic.
XX XX
XX Key Location/Qualifiers
XX FH repeat_unit 3..7
XX FT /*tag= a
XX FT repeat_unit 8..12
XX FT /*tag= b
XX FT repeat_unit 13..17
XX FT /*tag= c
XX FT repeat_unit 18..22
XX FT /*tag= d
XX XX
XX PN MO9616989-A1.
XX XX
XX PD 06-JUN-1996.
XX XX
XX PF 27-NOV-1995; 95WO-US015353.
XX XX
XX PR 28-NOV-1994; 94US-00347792.
XX PR 28-APR-1995; 95US-00431357.
XX PR 01-JUN-1995; 95US-00456623.
XX XX
XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX XX
XX PI Halazonecis TD;
XX XX
XX DR WPI; 1996-286828/29.
XX XX
XX PT New chimaeric p53 protein with heterologous tetramerisation domain - and
XX PT related DNA and vectors, useful for treating abnormal cell proliferation,
XX PT esp. cancer, auto-immune disease, etc.
XX XX
XX PS Example 1; Page 45; 123pp; English.
XX XX
XX CC Oligonucleotides BC.V4A (AAT32848) and BC (AAT32849) contain artificial
XX CC pentanucleotide repeat sites recognised by human tumour suppressor
XX CC protein p53 (see also AAM02617). They were used with oligonucleotide
XX CC Bxafi (AAT32847), which corresponds to the enhancer that drives p53-
XX CC dependent transcription of the waf1 gene, to assay p53-mediated
XX CC transcriptional activity and the DNA binding activity of wild-type p53
XX CC and of p53 proteins with altered tetramerisation domains
XX XX
XX SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
XX Query Match 84.0%; Score 16.8; DB 2; Length 30;
XX Best Local Similarity 90.0%; Pred. No. 2.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GGACATGCCCGGCGCATGTCC 20
Db 3 GGGCATGTCCGGCATGTCC 22

```

RESULT 36  
AAT32849/c  
ID AAT32849 standard; DNA; 30 BP.

XX AC AAT32849;  
XX XX  
XX DT 06-NOV-1996 (first entry)  
XX XX

```

DE DE Tumour suppressor p53 binding site oligonucleotide BC.
XX XX
XX p53 protein; tumour suppressor; tetramerisation domain;
XX KW chimaeric protein; GCN4; transcription factor; leucine zipper;
XX KW gene therapy; vector; cell proliferation; cancer; apoptosis;
XX KW autoimmune disease; immune tolerance; ds.
XX XX
XX OS Synthetic.
XX XX
XX Key Location/Qualifiers
XX FH repeat_unit 3..7
XX FT /*tag= a
XX FT repeat_unit 8..12
XX FT /*tag= b
XX FT repeat_unit 13..17
XX FT /*tag= c
XX FT repeat_unit 18..22
XX FT /*tag= d
XX XX
XX PN MO9616989-A1.
XX XX
XX PD 06-JUN-1996.
XX XX
XX PF 27-NOV-1995; 95WO-US015353.
XX XX
XX PR 28-NOV-1994; 94US-00347792.
XX PR 28-APR-1995; 95US-00431357.
XX PR 01-JUN-1995; 95US-00456623.
XX XX
XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX XX
XX PI Halazonecis TD;
XX XX
XX DR WPI; 1996-286828/29.
XX XX
XX PT New chimaeric p53 protein with heterologous tetramerisation domain - and
XX PT related DNA and vectors, useful for treating abnormal cell proliferation,
XX PT esp. cancer, auto-immune disease, etc.
XX XX
XX PS Example 1; Page 45; 123pp; English.
XX XX
XX CC Oligonucleotides BC.V4A (AAT32848) and BC (AAT32849) contain artificial
XX CC pentanucleotide repeat sites recognised by human tumour suppressor
XX CC protein p53 (see also AAM02617). They were used with oligonucleotide
XX CC Bxafi (AAT32847), which corresponds to the enhancer that drives p53-
XX CC dependent transcription of the waf1 gene, to assay p53-mediated
XX CC transcriptional activity and the DNA binding activity of wild-type p53
XX CC and of p53 proteins with altered tetramerisation domains
XX XX
XX SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
XX Query Match 84.0%; Score 16.8; DB 2; Length 30;
XX Best Local Similarity 90.0%; Pred. No. 2.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GGACATGCCCGGCGCATGTCC 20
Db 22 GGACATGCCCGGCGCATGTCC 3

```

RESULT 37  
AAT39778  
ID AAT39778 standard; DNA; 30 BP.

XX AC AAT39778;  
XX XX  
XX DT 30-APR-1997 (first entry)  
XX XX  
XX DE Competitor DNA oligonucleotide BC.  
XX DE  
XX KW Human; p53; cell proliferation; cell death; regulator; tumour; psoriasis;  
XX KW negative regulatory region; DNA damaging agent; transplant rejection;  
XX KW abnormal cell proliferation; atherosclerosis; cancer; autoimmune disease;

KW arterial restenosis; immune response; apoptosis; inducer; therapy;  
 KW proliferating lymphocytes; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9625434-A1.  
 XX  
 PD 22-AUG-1996.  
 XX  
 PF 16-FEB-1996; 96WO-US001535.  
 XX  
 PR 16-FEB-1995; 95US-00392542.  
 XX  
 PA (WIST-) WISTAR INST.  
 XX (FARB ) BAYER CORP.  
 XX  
 PI Halazonetis T, Hartwig W;  
 XX  
 DR WPI, 1996-393345/39.  
 XX  
 PT New human p53-isomorph peptide(s) and peptide:mimetic cpds. - used for  
 PT activating p53 function, e.g. for treating tumours, cancers, psoriasis,  
 PT etc.  
 PS Example 2; Page 24; 55pp; English.  
 XX  
 CC AA139774-139778 represent competitor DNA's for the peptides of the  
 CC invention (see AAM05350-W05374). The peptides of the invention are p53  
 CC fragments (see AAM05344 for full length wild type sequence). The p53  
 CC protein functions to regulate cell proliferation and cell death, and is  
 CC mutated in more than half of all human tumours. The peptide sequences are  
 CC used to activate the DNA binding activity of wild type p53, and p53  
 CC mutants (see AAM05345-W05349). The peptides of the invention consist of  
 CC at least four sequential amino acids from a negative regulatory region  
 CC which maps to residues 361-383 of p53. The peptide sequences preferably  
 CC contain four amino acids from a non-human p53 sequence, contain D-form  
 CC amino acids, and can also be cyclic peptides. The sequences retain the  
 CC structural characteristics of the original peptides, but the  
 CC modifications render them less susceptible to cleavage by proteases and  
 CC exopeptidases. As these sequences activate p53 DNA binding, they can be  
 CC used to identify p53 mutants. The peptides can also be used for treating  
 CC a patient with a tumour expressing a p53 mutant whose ability to bind DNA  
 CC may be activated by one of the peptides. They can also be used for  
 CC treating conditions such as exposure to DNA damaging agents, abnormal  
 CC cell proliferation characteristic of psoriasis, atherosclerosis, cancer,  
 CC arterial restenosis, autoimmune diseases and undesirable immune responses  
 CC accompanying rejection of a transplant. The peptides can also induce  
 CC apoptosis of specific cells, such as proliferating lymphocytes  
 CC  
 SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;  
 Query Match 84.0%; Score 16.8; DB 2; Length 30;  
 Best Local Similarity 90.0%; Pred. No. 2.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCGCATGCC 20  
 Db 3 GGCGATGTCGCGCGCATGCC 22  
 RESULT 38  
 AA139778/c  
 ID AA139778 standard; DNA, 30 BP.  
 XX  
 AC AA139778;  
 XX  
 DT 30-APR-1997 (first entry)  
 XX  
 DB Competitor DNA oligonucleotide BC.  
 XX  
 KW Human; p53; cell proliferation; cell death; regulator; tumour; psoriasis;  
 KW negative regulatory region; DNA damaging agent; transplant rejection;  
 KW abnormal cell proliferation; atherosclerosis; cancer; autoimmune disease;

KW arterial restenosis; immune response; apoptosis; inducer; therapy;  
 KW proliferating lymphocytes; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9625434-A1.  
 XX  
 PD 22-AUG-1996.  
 XX  
 PF 16-FEB-1996; 96WO-US001535.  
 XX  
 PR 16-FEB-1995; 95US-00392542.  
 XX  
 PA (WIST-) WISTAR INST.  
 XX (FARB ) BAYER CORP.  
 XX  
 PI Halazonetis T, Hartwig W;  
 XX  
 DR WPI, 1996-393345/39.  
 XX  
 PT New human p53-isomorph peptide(s) and peptide:mimetic cpds. - used for  
 PT activating p53 function, e.g. for treating tumours, cancers, psoriasis,  
 PT etc.  
 PS Example 2; Page 24; 55pp; English.  
 XX  
 CC AA139774-139778 represent competitor DNA's for the peptides of the  
 CC invention (see AAM05350-W05374). The peptides of the invention are p53  
 CC fragments (see AAM05344 for full length wild type sequence). The p53  
 CC protein functions to regulate cell proliferation and cell death, and is  
 CC mutated in more than half of all human tumours. The peptide sequences are  
 CC used to activate the DNA binding activity of wild type p53, and p53  
 CC mutants (see AAM05345-W05349). The peptides of the invention consist of  
 CC at least four sequential amino acids from a negative regulatory region  
 CC which maps to residues 361-383 of p53. The peptide sequences preferably  
 CC contain four amino acids from a non-human p53 sequence, contain D-form  
 CC amino acids, and can also be cyclic peptides. The sequences retain the  
 CC structural characteristics of the original peptides, but the  
 CC modifications render them less susceptible to cleavage by proteases and  
 CC exopeptidases. As these sequences activate p53 DNA binding, they can be  
 CC used to identify p53 mutants. The peptides can also be used for treating  
 CC a patient with a tumour expressing a p53 mutant whose ability to bind DNA  
 CC may be activated by one of the peptides. They can also be used for  
 CC treating conditions such as exposure to DNA damaging agents, abnormal  
 CC cell proliferation characteristic of psoriasis, atherosclerosis, cancer,  
 CC arterial restenosis, autoimmune diseases and undesirable immune responses  
 CC accompanying rejection of a transplant. The peptides can also induce  
 CC apoptosis of specific cells, such as proliferating lymphocytes  
 CC  
 SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;  
 Query Match 84.0%; Score 16.8; DB 2; Length 30;  
 Best Local Similarity 90.0%; Pred. No. 2.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCGCATGCC 20  
 Db 22 GGACATGCCCGGCGCATGCC 3  
 RESULT 39  
 AA160143  
 ID AA160143 standard; DNA, 30 BP.  
 XX  
 AC AA160143;  
 XX  
 DT 25-JUN-1997 (first entry)  
 XX  
 DB Oligonucleotide BC for testing DNA binding activity.  
 XX  
 KW p53; tumour suppressor; cancer; therapy; cell proliferation; apoptosis;  
 KW protein engineering; DNA binding; ds.

```

OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_binding 3..22
FT /tag= a
FT /note= "p53 DNA binding site"
XX
XX WO9710843-A1.
XX
XX 27-MAR-1997.
XX
XX 20-SEP-1996; 96WO-US015188.
XX
XX 22-SEP-1995; 95US-0004802P.
XX
XX 21-AUG-1996; 96US-00697221.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Halazonetis TD;
XX
XX WPI; 1997-202618/18.
XX
XX R284K modified p53 protein having DNA binding ability - useful in
XX treatment of cancer.
XX
XX Example 2; Page 20; 82pp; English.
XX
XX Oligonucleotide BC (AAT60143) contains a high affinity p53 binding site,
XX and was used with oligonucleotides Bp21 (AAT60144), Bgadd45 (AAT60145)
XX and non-specific TF3 (AAT60146) to examine the DNA binding activity of
XX p53 variants carrying a Thr284 to Arg substn. The results showed that the
XX T284R substn. enhanced binding to all p53 proteins examined. For wild-
XX type p53 (see also AAW13949), the effect was evident with
XX oligonucleotides BC and Bp21, for p53Gln248 (see also AAW13968) it was
XX evident for oligonucleotide BC, and for p53His273 (see also AAW13969) and
XX p53Cys273 (see also AAW13970) it was evident for oligonucleotides BC,
XX Bp21 and Bgadd45
XX
XX Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
SQ
XX
XX Query Match 84.0%; Score 16.8; DB 2; Length 30;
XX Best Local Similarity 90.0%; Pred. No. 2.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 GGACATGCCCGGCGCATGTC 20
XX 3 GGCGATGTCGCGGCGCATGTC 22
XX
XX Db
XX
XX RESULT 40
XX AAT60143/c
XX ID AAT60143 standard; DNA; 30 BP.
XX
XX AAT60143;
XX
XX 25-JUN-1997 (first entry)
XX
XX Oligonucleotide BC for testing DNA binding activity.
XX
XX p53; tumour suppressor; cancer; therapy; cell proliferation; apoptosis;
XX protein engineering; DNA binding; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH misc_binding 3..22
XX FT /tag= a
XX FT /note= "p53 DNA binding site"
XX
XX WO9710843-A1.
XX
XX 27-MAR-1997.
XX

```

```

PF 20-SEP-1996; 96WO-US015188.
XX
XX 22-SEP-1995; 95US-0004802P.
XX
XX 21-AUG-1996; 96US-00697221.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Halazonetis TD;
XX
XX WPI; 1997-202618/18.
XX
XX R284K modified p53 protein having DNA binding ability - useful in
XX treatment of cancer.
XX
XX Example 2; Page 20; 82pp; English.
XX
XX Oligonucleotide BC (AAT60143) contains a high affinity p53 binding site,
XX and was used with oligonucleotides Bp21 (AAT60144), Bgadd45 (AAT60145)
XX and non-specific TF3 (AAT60146) to examine the DNA binding activity of
XX p53 variants carrying a Thr284 to Arg substn. The results showed that the
XX T284R substn. enhanced binding to all p53 proteins examined. For wild-
XX type p53 (see also AAW13949), the effect was evident with
XX oligonucleotides BC and Bp21, for p53Gln248 (see also AAW13968) it was
XX evident for oligonucleotide BC, and for p53His273 (see also AAW13969) and
XX p53Cys273 (see also AAW13970) it was evident for oligonucleotides BC,
XX Bp21 and Bgadd45
XX
XX Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
SQ
XX
XX Query Match 84.0%; Score 16.8; DB 2; Length 30;
XX Best Local Similarity 90.0%; Pred. No. 2.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 GGACATGCCCGGCGCATGTC 20
XX 22 GGACATGCCCGGCGCATGTC 3
XX
XX Db
XX
XX RESULT 41
XX ADU60926
XX ID ADU60926 standard; DNA; 4157 BP.
XX
XX ADU60926;
XX
XX 06-MAY-2004 (first entry)
XX
XX Concatemer of PDB4A oligonucleotides.
XX
XX interleukin; IL-4 receptor; IL-5 receptor; lung disease;
XX airway inflammation; allergy; asthma; impeded respiration;
XX cystic fibrosis; acute respiratory distress syndrome;
XX pulmonary hypertension; lung inflammation; bronchitis; ds.
XX
XX Homo sapiens.
XX
XX WO2004011613-A2.
XX
XX 05-FEB-2004.
XX
XX 25-JUL-2003; 2003WO-US023509.
XX
XX 29-JUL-2002; 2002US-0399076P.
XX
XX (EPIC-) EPICGENESIS PHARM INC.
XX
XX Nyce JW, Tang L, Sandrasagra A, Aguilar D, Miller S;
XX Shahbuddin S, Lu H, Cong H;
XX WPI; 2004-203534/19.
XX
XX Novel single or multiple target oligonucleotide anti-sense to e.g.
XX PT initiation codons and introns of respiratory disease-relevant genes e.g.,
XX CCR1, RANTES, MCP4, useful for prophylaxis or treating respiratory
XX

```

PT disease e.g., asthma.  
 XX  
 PS Claim 2; SEQ ID NO 1782; 85pp; English.  
 XX  
 CC The present invention relates to an oligonucleotide anti-sense to e.g.,  
 CC interleukin (IL)-4 receptor, IL-5 receptor or salts of the  
 CC oligonucleotide. The method is useful for preventing or treating a  
 CC respiratory or lung disease, which involves administering to the airways  
 CC of a subject an effective amount of an inhibitor. The oligonucleotide is  
 CC useful for production of a medicament for the prevention and/or treatment  
 CC of a respiratory or lung disease. The respiratory or lung disease is  
 CC chosen from allergy inflammation, allergy (ies), asthma, impeded  
 CC respiration, cystic fibrosis (CF), chronic obstructive pulmonary diseases  
 CC (COPD), allergic rhinitis (AR), acute respiratory distress syndrome  
 CC (ARDS), pulmonary hypertension, lung inflammation, bronchitis, airway  
 CC obstruction. The present sequence represents a concatenation of  
 CC oligonucleotides of the invention.  
 XX  
 SQ Sequence 4157 BP; 750 A; 1181 C; 1250 G; 976 T; 0 U; 0 Other;  
 Query Match 84.0%; Score 16.8; DB 12; Length 4157;  
 Best Local Similarity 90.0%; Pred. No. 2.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGACATGTC 20  
 DB 2031 GGACATGTCGACATGTC 2050  
 RESULT 42  
 RES0926/C  
 ID AD060926 standard; DNA; 4157 BP.  
 XX  
 AC AD060926;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Concatenation of PDBA4 oligonucleotides.  
 XX  
 KW Interleukin; IL-4 receptor; IL-5 receptor; lung disease;  
 KW airway inflammation; allergy; asthma; impeded respiration;  
 KW cystic fibrosis; acute respiratory distress syndrome;  
 KW pulmonary hypertension; lung inflammation; bronchitis; ds.  
 KW  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 PN WO2004011613-A2.  
 XX  
 PD 05-FEB-2004.  
 XX  
 PF 25-JUL-2003; 2003WO-US023509.  
 XX  
 PR 29-JUL-2002; 2002US-0399076P.  
 XX  
 PA (EPIC-) EPICGENESIS PHARM INC.  
 XX  
 PI NYCE JM, Tang L, Sandrasagra A, Aguilar D, Miller S;  
 PI Shahabuddin S, Lu H, Cong H;  
 XX  
 DR WPI; 2004-203534/19.  
 XX  
 PT Novel single or multiple target oligonucleotide anti-sense to e.g.  
 PT interleukin codons and introns of respiratory disease-relevant genes e.g.,  
 PT CCR1, RANTES, MCP4, useful for prophylaxis or treating respiratory  
 PT disease e.g., asthma.  
 XX  
 PS Claim 2; SEQ ID NO 1782; 85pp; English.  
 CC  
 CC The present invention relates to an oligonucleotide anti-sense to e.g.,  
 CC interleukin codon, coding region with 2-10 nucleotides of 5'-end and 3'-

CC end of nucleic acid target comprising gene(s) chosen from e.g.  
 CC interleukin (IL)-4 receptor, IL-5 receptor or salts of the  
 CC oligonucleotide. The method is useful for preventing or treating a  
 CC respiratory or lung disease, which involves administering to the airways  
 CC of a subject an effective amount of an inhibitor. The oligonucleotide is  
 CC useful for production of a medicament for the prevention and/or treatment  
 CC of a respiratory or lung disease. The respiratory or lung disease is  
 CC chosen from allergy inflammation, allergy (ies), asthma, impeded  
 CC respiration, cystic fibrosis (CF), chronic obstructive pulmonary diseases  
 CC (COPD), allergic rhinitis (AR), acute respiratory distress syndrome  
 CC (ARDS), pulmonary hypertension, lung inflammation, bronchitis, airway  
 CC obstruction. The present sequence represents a concatenation of  
 CC oligonucleotides of the invention.  
 XX  
 SQ Sequence 4157 BP; 750 A; 1181 C; 1250 G; 976 T; 0 U; 0 Other;  
 Query Match 84.0%; Score 16.8; DB 12; Length 4157;  
 Best Local Similarity 90.0%; Pred. No. 2.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGACATGTC 20  
 DB 2050 GGACATGTCGACATGTC 2031  
 RESULT 43  
 ADO46415  
 ID ADO46415 standard; DNA; 4167 BP.  
 XX  
 AC ADO46415;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Human oligonucleotide #1781.  
 XX  
 KW Human; ss; interleukin-4 receptor; IL-4; interleukin-5 receptor; IL-5;  
 KW CCR1; CCR3; Botaxin-1; RANTES; MCP4; CD23; ICAM; VCAM; tryptase a;  
 KW tryptase b; PDBA 4; PDBA B; PDBA C; PDBA D; respiratory disease;  
 KW lung disease; hyper-responsiveness; adenosine; adenosine A receptor;  
 KW asthma; lung allergy; inflammation; inflammatory disease;  
 KW airway inflammation; allergy; impeded respiration; cystic fibrosis; CF;  
 KW chronic obstructive pulmonary disease; COPD; allergic rhinitis;  
 KW acute respiratory distress syndrome; pulmonary hypertension;  
 KW lung inflammation; bronchitis; airway obstruction; bronchoconstriction.  
 KW  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 PN US2004049022-A1.  
 XX  
 PD 11-MAR-2004.  
 XX  
 PF 25-JUL-2003; 2003US-00627930.  
 XX  
 PR 23-APR-2002; 2002WO-US013135.  
 XX  
 PR 23-APR-2002; 2002WO-US013143.  
 XX  
 PA (NYCE/) NYCE J W.  
 PA (SAND/) SANDRASAGRA A.  
 PA (TANG/) TANG L.  
 PA (AGUI/) AGUILAR D.  
 PA (MILL/) MILLER S.  
 PA (SHAH/) SHAHABUDDIN S.  
 PA (LUH/) LU H.  
 PA (CONG/) CONG H.  
 XX  
 PI NYCE JM, Sandrasagra A, Tang L, Aguilar D, Miller S;  
 PI Shahabuddin S, Lu H, Cong H;  
 XX  
 DR WPI; 2004-293804/27.  
 XX  
 PT Novel single or multiple target oligonucleotide anti-sense to e.g.  
 PT interleukin codon, intron of respiratory disease-relevant gene e.g. CCR1,  
 PT interleukin codon, coding region with 2-10 nucleotides of 5'-end and 3'-

PT RANTES, MCP4, useful for prophylaxis or treating respiratory disease e.g. asthma.

Claim 2; SEQ ID NO 1782; 174pp; English.

The invention relates to oligonucleotides anti-sense to an initiation codon, coding region, 5' or 3' intron-exon junction, intron or region with 2-10 nucleotides of the 5'-end or 3'-end of a nucleic acid target chosen from a gene encoding interleukin (IL)-4 receptor, interleukin (IL)-5 receptor, CCR1, CCR3, Eotaxin-1, RANTES, MCP4, CD23, ICAM, VCAM, tryptase a, tryptase b, PDB4 A, PDB4 B, PDB4 C or PDB4 D. The invention also relates to a method of screening a candidate compound that binds to one or more nucleic acid target(s) or expressed product(s), for the prevention and/or treatment of a respiratory or lung disease. The oligonucleotides are useful for reducing or inhibiting expression of a gene or mRNA encoding interleukin-4 receptor, interleukin-5 receptor, CCR1, CCR3, Eotaxin-1, RANTES, MCP4, CD23, ICAM, VCAM, tryptase a, tryptase b, PDB4 A, PDB4 B, PDB4 C, or PDB4 D. The oligonucleotides are useful for preventing or treating a respiratory or lung disease. The receptor(s) and/or asthma and/or lung allergies associated with inflammation or an inflammatory disease. The respiratory or lung disease is chosen from airway inflammation, allergy, asthma, impeded respiration, cystic fibrosis (CF), chronic obstructive pulmonary disease (COPD), allergic rhinitis, acute respiratory distress syndrome, pulmonary hypertension, lung inflammation, bronchitis, airway obstruction or bronchoconstriction. This sequence represents an oligonucleotide of the invention.

Sequence 4167 BP; 750 A; 1185 C; 1250 G; 982 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 12; Length 4167;  
Best Local Similarity 90.0%; Pred. NO. 2.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
Db 2041 GGACATGTCCGGACATGTCC 2060

RESULT 44  
ADO46415/c  
ID ADO46415 standard; DNA; 4167 BP.

XX ADO46415;

XX 15-JUL-2004 (first entry)

XX Human oligonucleotide #1781.

XX Human; ss; interleukin-4 receptor; IL-4; interleukin-5 receptor; IL-5; CCR1; CCR3; Eotaxin-1; RANTES; MCP4; CD23; ICAM; VCAM; tryptase a; tryptase b; PDB4 A; PDB4 B; PDB4 C; PDB4 D; respiratory disease; lung disease; hyper-responsiveness; adenosine; adenosine A receptor; asthma; lung allergy; inflammation; inflammatory disease; airway inflammation; allergy; impeded respiration; cystic fibrosis; CF; chronic obstructive pulmonary disease; COPD; allergic rhinitis; acute respiratory distress syndrome; pulmonary hypertension; lung inflammation; bronchitis; airway obstruction; bronchoconstriction.

XX Homo sapiens.

XX US2004049022-A1.

XX 11-MAR-2004.

XX 25-JUL-2003; 2003US-00627930.

XX 23-APR-2002; 2002WO-US013135.

XX 23-APR-2002; 2002WO-US013143.

XX (NYCE/) NYCE J W.

PA (SAND/) SANDRASAGRA A.

PA (TANG/) TANG L.

PA (AGUI/) AGUIAR D.

PA (MILL/) MILLER S.

PA (SHAH/) SHAHABUDDIN S.

PA (LUHH/) LU H.

PA (CONG/) CONG H.

PI Nye JW, Sandrasagra A, Tang L, Aguiar D, Miller S;

PI Shahabuddin S, Lu H, Cong H;

PI WPI; 2004-293804/27.

PT Novel single or multiple target oligonucleotide anti-sense to e.g.

PT initiation codon, intron of respiratory disease-relevant gene e.g. CCR1,

PT RANTES, MCP4, useful for prophylaxis or treating respiratory disease e.g.

PT asthma.

PT Claim 2; SEQ ID NO 1782; 174pp; English.

The invention relates to oligonucleotides anti-sense to an initiation codon, coding region, 5' or 3' intron-exon junction, intron or region with 2-10 nucleotides of the 5'-end or 3'-end of a nucleic acid target chosen from a gene encoding interleukin (IL)-4 receptor, interleukin (IL)-5 receptor, CCR1, CCR3, Eotaxin-1, RANTES, MCP4, CD23, ICAM, VCAM, tryptase a, tryptase b, PDB4 A, PDB4 B, PDB4 C or PDB4 D. The invention also relates to a method of screening a candidate compound that binds to one or more nucleic acid target(s) or expressed product(s), for the prevention and/or treatment of a respiratory or lung disease. The oligonucleotides are useful for reducing or inhibiting expression of a gene or mRNA encoding interleukin-4 receptor, interleukin-5 receptor, CCR1, CCR3, Eotaxin-1, RANTES, MCP4, CD23, ICAM, VCAM, tryptase a, tryptase b, PDB4 A, PDB4 B, PDB4 C, or PDB4 D. The oligonucleotides are useful for preventing or treating a respiratory or lung disease. The receptor(s) and/or asthma and/or lung allergies associated with inflammation or an inflammatory disease. The respiratory or lung disease is chosen from airway inflammation, allergy, asthma, impeded respiration, cystic fibrosis (CF), chronic obstructive pulmonary disease (COPD), allergic rhinitis, acute respiratory distress syndrome, pulmonary hypertension, lung inflammation, bronchitis, airway obstruction or bronchoconstriction. This sequence represents an oligonucleotide of the invention.

SO Sequence 4167 BP; 750 A; 1185 C; 1250 G; 982 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 12; Length 4167;  
Best Local Similarity 90.0%; Pred. NO. 2.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
Db 2060 GGACATGTCCGGACATGTCC 2041

RESULT 45

AA88896  
ID AA88896 standard; cDNA; 9391 BP.

XX AA88896;

XX 05-MAR-2001 (first entry)

XX Mouse laminin 2 alpha-2 chain cDNA.

XX Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;

XX degenerative muscle disorder; muscular dystrophy; cell therapy; ss.

XX Mus musculus.

XX Key mat\_peptide Location/Qualifiers

XX 1.9252



```

PT      /tag= a
XX      WO200066730-A2.
PN      09-NOV-2000.
XX      28-APR-2000; 2000WO-US011378.
XX      30-APR-1999; 99US-0131720P.
XX      15-JUN-1999; 99US-0139198P.
XX      12-JUL-1999; 99US-0143289P.
XX      24-SEP-1999; 99US-0155945P.
XX      (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX      Yurchenco P;
XX      WPI; 2000-687537/67.
XX      P-PSDB; AAB19796.
XX      Purified laminin 2 protein, useful for research and therapeutic purposes
PT      including peripheral nerve regeneration, treatment of degenerative muscle
PT      disorders, angiogenesis regulation, and ex vivo cell therapy.
XX      Claim 4; Page 156-168; 305pp; English.
XX      The present sequence is that of cDNA encoding the mature alpha-2 chain
XX      (see AAB19796) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400
XX      kDa), beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX      specifically required for stabilizing myotubes during skeletal muscle
XX      development, and for preventing apoptosis. Genetic defects in human
XX      laminin 2 structure or expression are associated with a major type of
XX      congenital muscular dystrophy. Laminin 2 is also thought to be important
XX      in Schwann cell/basal lamina interactions. The invention provides laminin
XX      2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX      the polynucleotides encoding them (see AAB88891-906), methods for making
XX      recombinant laminin 2, cells that express recombinant laminin 2, and
XX      methods for using purified laminin 2 for research and therapeutic
XX      purposes including peripheral nerve regeneration, treatment of
XX      degenerative muscle disorders, angiogenesis regulation, promoting cell
XX      attachment and migration, ex vivo cell therapy, improving the take of
XX      grafts, improving the biocompatibility of medical devices and preparing
XX      improved culture devices and media
XX      Sequence 9391 BP; 2653 A; 2153 C; 2334 G; 2241 T; 0 U; 0 Other;
SQ
Query Match      84.0%; Score 16.8; DB 3; Length 9391;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 GGACATGCCCGGCGATGCC 20
DB      793 GGTGATGCCCGGCTTGTCC 812

```

RESULT 46  
 ID AAB88896 standard; CDNA; 9391 BP.

AC AAA88896;  
 XX  
 DT 05-MAR-2001 (first entry)  
 XX  
 DB Mouse laminin 2 alpha-2 chain cDNA.  
 XX  
 KW Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;  
 KW degenerative muscle disorder; muscular dystrophy; cell therapy; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT mat\_peptide 1..9252  
 FT /tag= a

```

XX      WO200066730-A2.
PN      09-NOV-2000.
XX      28-APR-2000; 2000WO-US011378.
XX      30-APR-1999; 99US-0131720P.
XX      15-JUN-1999; 99US-0139198P.
XX      12-JUL-1999; 99US-0143289P.
XX      24-SEP-1999; 99US-0155945P.
XX      (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX      Yurchenco P;
XX      WPI; 2000-687537/67.
XX      P-PSDB; AAB19796.
XX      Purified laminin 2 protein, useful for research and therapeutic purposes
PT      including peripheral nerve regeneration, treatment of degenerative muscle
PT      disorders, angiogenesis regulation, and ex vivo cell therapy.
XX      Claim 4; Page 156-168; 305pp; English.
XX      The present sequence is that of cDNA encoding the mature alpha-2 chain
XX      (see AAB19796) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400
XX      kDa), beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX      specifically required for stabilizing myotubes during skeletal muscle
XX      development, and for preventing apoptosis. Genetic defects in human
XX      laminin 2 structure or expression are associated with a major type of
XX      congenital muscular dystrophy. Laminin 2 is also thought to be important
XX      in Schwann cell/basal lamina interactions. The invention provides laminin
XX      2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX      the polynucleotides encoding them (see AAB88891-906), methods for making
XX      recombinant laminin 2, cells that express recombinant laminin 2, and
XX      methods for using purified laminin 2 for research and therapeutic
XX      purposes including peripheral nerve regeneration, treatment of
XX      degenerative muscle disorders, angiogenesis regulation, promoting cell
XX      attachment and migration, ex vivo cell therapy, improving the take of
XX      grafts, improving the biocompatibility of medical devices and preparing
XX      improved culture devices and media
XX      Sequence 9391 BP; 2663 A; 2153 C; 2334 G; 2241 T; 0 U; 0 Other;
SQ
Query Match      84.0%; Score 16.8; DB 3; Length 9391;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 GGACATGCCCGGCGATGCC 20
DB      812 GGACATGCCCGGCGATGCC 793

```

RESULT 47  
 ID AAB88895 standard; CDNA; 9511 BP.

AC AAA88895;  
 XX  
 DT 05-MAR-2001 (first entry)  
 XX  
 DB Mouse laminin 2 alpha-2 chain cDNA.  
 XX  
 KW Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;  
 KW degenerative muscle disorder; muscular dystrophy; cell therapy; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 55..9375  
 FT /tag= a  
 FT sig\_peptide 55..120

```

FT mat_peptide      /*tag= b
FT 121..9372
FT /*tag= c
XX
XX MO200066730-A2.
XX
XX 09-NOV-2000.
XX
XX 28-APR-2000; 2000WO-US011378.
XX
XX 30-APR-1999; 99US-0131720P.
XX 15-JUN-1999; 99US-0139198P.
XX 12-JUL-1999; 99US-0143289P.
XX 24-SEP-1999; 99US-0155945P.
XX
XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
XX Yurchenco P;
XX
XX WPI; 2000-687537/67.
XX P-PSDB; AAB19795.
XX
XX Purified laminin 2 protein, useful for research and therapeutic purposes
XX including peripheral nerve regeneration, treatment of degenerative muscle
XX disorders, angiogenesis regulation, and ex vivo cell therapy.
XX
XX Claim 4; Page 133-146; 305pp; English.
XX
XX The present sequence is that of cDNA encoding the alpha-2 chain (see
XX AAB19795) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400 kDa),
XX beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX specifically required for stabilizing myotubes during skeletal muscle
XX development, and for preventing apoptosis. Genetic defects in its
XX structure or expression are associated with a major type of congenital
XX muscular dystrophy in humans. Laminin 2 is also thought to be important
XX in Schwann cell/basal lamina interactions. The invention provides laminin
XX 2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX recombinant laminin 2, cells that express recombinant laminin 2, and
XX methods for using purified laminin 2 for research and therapeutic
XX purposes including peripheral nerve regeneration, treatment of
XX degenerative muscle disorders, angiogenesis regulation, promoting cell
XX attachment and migration, ex vivo cell therapy, improving the take of
XX grafts, improving the biocompatibility of medical devices and preparing
XX improved culture devices and media
XX
XX Sequence 9511 BP; 2679 A; 2194 C; 2375 G; 2263 T; 0 U; 0 Other;
XX
XX Query Match      84.0%; Score 16.8; DB 3; Length 9511;
XX Best Local Similarity 90.0%; Pred. No. 2.5e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GGACATGCCCGGCGCATGTCC 20
Db 913 GGTCATGCCCGGCGCTGTCC 932

```

RESULT 48  
AAA88895/c  
ID AAA88895 standard; cDNA; 9511 BP.  
AC AAA88895;  
XX  
DT 05-MAR-2001 (first entry)  
XX  
DE Mouse laminin 2 alpha-2 chain cDNA.  
XX  
KM Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;  
XX degenerative muscle disorder; muscular dystrophy; cell therapy; ss.  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers

```

FT CDS      55..9375
FT sig_peptide /*tag= a
FT 55..120
FT /*tag= b
FT mat_peptide /*tag= c
FT 121..9372
XX
XX MO200066730-A2.
XX
XX 09-NOV-2000.
XX
XX 28-APR-2000; 2000WO-US011378.
XX
XX 30-APR-1999; 99US-0131720P.
XX 15-JUN-1999; 99US-0139198P.
XX 12-JUL-1999; 99US-0143289P.
XX 24-SEP-1999; 99US-0155945P.
XX
XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
XX Yurchenco P;
XX
XX WPI; 2000-687537/67.
XX P-PSDB; AAB19795.
XX
XX Purified laminin 2 protein, useful for research and therapeutic purposes
XX including peripheral nerve regeneration, treatment of degenerative muscle
XX disorders, angiogenesis regulation, and ex vivo cell therapy.
XX
XX Claim 4; Page 133-146; 305pp; English.
XX
XX The present sequence is that of cDNA encoding the alpha-2 chain (see
XX AAB19795) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400 kDa),
XX beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX specifically required for stabilizing myotubes during skeletal muscle
XX development, and for preventing apoptosis. Genetic defects in its
XX structure or expression are associated with a major type of congenital
XX muscular dystrophy in humans. Laminin 2 is also thought to be important
XX in Schwann cell/basal lamina interactions. The invention provides laminin
XX 2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX recombinant laminin 2, cells that express recombinant laminin 2, and
XX methods for using purified laminin 2 for research and therapeutic
XX purposes including peripheral nerve regeneration, treatment of
XX degenerative muscle disorders, angiogenesis regulation, promoting cell
XX attachment and migration, ex vivo cell therapy, improving the take of
XX grafts, improving the biocompatibility of medical devices and preparing
XX improved culture devices and media
XX
XX Sequence 9511 BP; 2679 A; 2194 C; 2375 G; 2263 T; 0 U; 0 Other;
XX
XX Query Match      84.0%; Score 16.8; DB 3; Length 9511;
XX Best Local Similarity 90.0%; Pred. No. 2.5e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GGACATGCCCGGCGCATGTCC 20
Db 932 GGACAAAGCCCGGCGCATGACC 913

```

RESULT 49  
ADA02555  
ID ADA02555 standard; DNA; 21565 BP.  
AC ADA02555;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Mouse lymphoma associated gene, SEQ ID NO:1073.  
XX  
KM Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;  
XX prostate; lymphoma; leukemia; cytostatic; gene therapy; drug screening;  
XX gene; ds.  
KM

```

XX  Mus sp.
OS
XX
XX  WO2003057146-A2.
PN
XX
XX  17-JUL-2003.
PD
XX
XX  26-DEC-2002; 2002WO-US041414.
PF
XX
XX  26-DEC-2001; 2001US-00035832.
PR
XX
XX  (SAGR-) SAGRES DISCOVERY.
PA
XX
XX  Morris DW;
PI
XX
XX  WPI; 2003-587068/55.
DR
XX
XX  New recombinant nucleic acid encoding carcinoma associated protein,
PT  useful for preparing compositions for treating carcinomas.
PS  Claim 1; SEQ ID NO 1073; 245bp; English.
XX
XX  The invention relates to recombinant carcinoma associated (CA) nucleic
CC  acid sequences from mouse and human (ADA01482-ADA03094), and to
CC  recombinant carcinoma associated proteins (CAP) encoded by them. The
CC  invention also encompasses expression vectors and host cells comprising a
CC  CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC  binds to the protein, and a biochip comprising CA nucleic acid or
CC  fragments thereof. The sequences of the invention were identified using
CC  oncogenic retroviruses, which insert into the genome of the host organism
CC  at random. Many of these do not carry transduced host oncogenes or
CC  pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC  direct consequence of the effects of proviral integration into host
CC  protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC  carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC  leukaemia) or a propensity to carcinoma by determination of the sequence
CC  of a CA gene, or by determination of CA gene expression in particular
CC  tissues. CA nucleic acids, proteins and antibodies are also useful as
CC  therapeutic agents and in screening and evaluating drug candidates. The
CC  present sequence represents a specifically claimed murine CA nucleic acid
CC  sequence of the invention. Note: The complete sequence data for this
CC  patent did not form part of the printed specification, but was obtained
CC  in electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 21565 BP; 5397 A; 5271 C; 5466 G; 5371 T; 0 U; 60 Other;
Query Match      84.0%; Score 16.8; DB 9; Length 21565;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY  1 GGACATGCCCGGCGCATGTCC 20
DB  9378 GGACATGCCCGGCGCATGTCC 9378

```

```

RESULT 50
ADA02555/C
ID  ADA02555 standard; DNA; 21565 BP.
XX
XX  ADA02555;
AC
XX
XX  06-NOV-2003 (first entry)
DT
XX
XX  Mouse lyfe carcinoma associated gene, SEQ ID NO:1073.
DE
XX
XX  Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KM  prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW  gene; ds.
XX
XX  Mus sp.
OS
XX
XX  WO2003057146-A2.
PN

```

```

XX  17-JUL-2003.
PD
XX
XX  26-DEC-2002; 2002WO-US041414.
PF
XX
XX  26-DEC-2001; 2001US-00035832.
PR
XX
XX  (SAGR-) SAGRES DISCOVERY.
PA
XX
XX  Morris DW;
PI
XX
XX  WPI; 2003-587068/55.
DR
XX
XX  New recombinant nucleic acid encoding carcinoma associated protein,
PT  useful for preparing compositions for treating carcinomas.
PS  Claim 1; SEQ ID NO 1073; 245bp; English.
XX
XX  The invention relates to recombinant carcinoma associated (CA) nucleic
CC  acid sequences from mouse and human (ADA01482-ADA03094), and to
CC  recombinant carcinoma associated proteins (CAP) encoded by them. The
CC  invention also encompasses expression vectors and host cells comprising a
CC  CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC  binds to the protein, and a biochip comprising CA nucleic acid or
CC  fragments thereof. The sequences of the invention were identified using
CC  oncogenic retroviruses, which insert into the genome of the host organism
CC  at random. Many of these do not carry transduced host oncogenes or
CC  pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC  direct consequence of the effects of proviral integration into host
CC  protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC  carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC  leukaemia) or a propensity to carcinoma by determination of the sequence
CC  of a CA gene, or by determination of CA gene expression in particular
CC  tissues. CA nucleic acids, proteins and antibodies are also useful as
CC  therapeutic agents and in screening and evaluating drug candidates. The
CC  present sequence represents a specifically claimed murine CA nucleic acid
CC  sequence of the invention. Note: The complete sequence data for this
CC  patent did not form part of the printed specification, but was obtained
CC  in electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 21565 BP; 5397 A; 5271 C; 5466 G; 5371 T; 0 U; 60 Other;
Query Match      84.0%; Score 16.8; DB 9; Length 21565;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY  1 GGACATGCCCGGCGCATGTCC 20
DB  9378 GGACATGCCCGGCGCATGTCC 9378

```

Search completed: January 14, 2005, 16:14:16  
Job time : 360.737 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:47:55 / Search time 71.0526 Seconds  
(without alignments)  
200.074 Million cell updates/sec

Title: US-09-578-453-2  
Perfect score: 20  
Sequence: 1 GCACATCCCGCGCATGTC 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 500 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	2	US-08-904-901-162
2	20	100.0	20	2	US-08-904-901-162
3	20	100.0	20	2	US-08-904-901-162
4	20	100.0	20	3	US-09-249-730-162
5	20	100.0	20	4	US-09-249-730-162
6	20	100.0	20	4	US-09-249-247-162
7	20	100.0	20	4	US-09-249-247-162
8	20	100.0	26	3	US-08-767-942A-40
9	20	100.0	27	3	US-08-767-942A-40
10	20	100.0	27	3	US-08-767-942A-41
11	20	100.0	30	1	US-08-347-792-18
12	20	100.0	30	1	US-08-347-792-18
13	20	100.0	30	1	US-08-431-357-18
14	20	100.0	30	1	US-08-431-357-18
15	20	100.0	30	1	US-08-431-357-18
16	20	100.0	30	2	US-08-697-221-29
17	20	100.0	30	2	US-08-697-221-29
18	20	100.0	30	3	US-08-392-542-31
19	20	100.0	30	3	US-08-392-542-31
20	20	100.0	30	3	US-08-894-327-31
21	20	100.0	30	4	US-09-685-027-31
22	20	100.0	30	4	US-09-685-027-31
23	20	100.0	30	4	US-09-829-922-31
24	20	100.0	30	4	US-09-829-922-31
25	20	100.0	30	4	US-09-829-922-31
26	20	100.0	30	5	PCT-US95-15353-18
27	20	100.0	30	5	PCT-US95-15353-18

C 28	16.8	84.0	9391	4	US-09-562-702A-11	Sequence 11, Appl
C 29	16.8	84.0	9511	4	US-09-562-702A-9	Sequence 9, Appl
C 30	16.8	84.0	9511	4	US-09-562-702A-9	Sequence 9, Appl
C 31	15.4	77.0	473	4	US-09-312-283C-355	Sequence 355, App
C 32	15.4	77.0	473	4	US-09-312-283C-355	Sequence 355, App
C 33	15.2	76.0	20	2	US-08-657-828A-3	Sequence 3, Appl
C 34	15.2	76.0	20	2	US-08-657-828A-3	Sequence 3, Appl
C 35	15.2	76.0	20	3	US-09-260-440-3	Sequence 3, Appl
C 36	15.2	76.0	20	3	US-09-260-440-3	Sequence 3, Appl
C 37	15.2	76.0	291	4	US-09-513-999C-22333	Sequence 22333, A
C 38	15.2	76.0	291	4	US-09-513-999C-22333	Sequence 22333, A
C 39	15.2	76.0	1251	4	US-09-828-000-1	Sequence 1, Appl
C 40	15.2	76.0	1251	4	US-09-828-000-1	Sequence 1, Appl
C 41	15.2	76.0	1920	4	US-09-906-393A-35	Sequence 35, Appl
C 42	15.2	76.0	1920	4	US-09-906-393A-35	Sequence 35, Appl
C 43	15.2	76.0	1958	4	US-09-702-327-3	Sequence 3, Appl
C 44	15.2	76.0	1958	4	US-09-702-327-3	Sequence 3, Appl
C 45	15.2	76.0	6942	2	US-08-460-309-3	Sequence 3, Appl
C 46	15.2	76.0	6942	2	US-08-460-309-3	Sequence 3, Appl
C 47	15.2	76.0	6942	2	US-08-125-077-3	Sequence 3, Appl
C 48	15.2	76.0	6942	2	US-08-125-077-3	Sequence 3, Appl
C 49	15.2	76.0	9419	4	US-09-562-702A-7	Sequence 7, Appl
C 50	15.2	76.0	9419	4	US-09-562-702A-7	Sequence 7, Appl
C 51	15.2	76.0	9420	4	US-09-562-702A-3	Sequence 3, Appl
C 52	15.2	76.0	9420	4	US-09-562-702A-3	Sequence 3, Appl
C 53	15.2	76.0	9534	4	US-09-562-702A-5	Sequence 5, Appl
C 54	15.2	76.0	9534	4	US-09-562-702A-5	Sequence 5, Appl
C 55	15.2	76.0	9534	4	US-09-561-709B-8	Sequence 8, Appl
C 56	15.2	76.0	9534	4	US-09-561-709B-8	Sequence 8, Appl
C 57	15.2	76.0	9534	4	US-09-917-254-35	Sequence 35, Appl
C 58	15.2	76.0	9534	4	US-09-917-254-35	Sequence 35, Appl
C 59	15.2	76.0	9535	4	US-09-562-702A-1	Sequence 1, Appl
C 60	15.2	76.0	9535	4	US-09-562-702A-1	Sequence 1, Appl
C 61	15.2	76.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 62	15.2	76.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 63	15.2	76.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C 64	15.2	76.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C 65	15	75.0	1743	4	US-09-252-991A-13408	Sequence 13408, A
C 66	15	75.0	1743	4	US-09-252-991A-13408	Sequence 13408, A
C 67	15	75.0	2199	4	US-09-252-991A-13503	Sequence 13503, A
C 68	15	75.0	2199	4	US-09-252-991A-13503	Sequence 13503, A
C 69	15	75.0	2433	4	US-09-252-991A-13950	Sequence 13950, A
C 70	15	75.0	2433	4	US-09-252-991A-13950	Sequence 13950, A
C 71	14.8	74.0	49	2	US-08-299-074A-31	Sequence 31, Appl
C 72	14.8	74.0	49	2	US-08-299-074A-31	Sequence 31, Appl
C 73	14.8	74.0	49	3	US-09-399-773-31	Sequence 31, Appl
C 74	14.8	74.0	49	3	US-09-399-773-31	Sequence 31, Appl
C 75	14.8	74.0	516	4	US-09-252-991A-4742	Sequence 4742, Ap
C 76	14.8	74.0	516	4	US-09-252-991A-4742	Sequence 4742, Ap
C 77	14.8	74.0	762	4	US-09-252-991A-4766	Sequence 4766, Ap
C 78	14.8	74.0	762	4	US-09-252-991A-4766	Sequence 4766, Ap
C 79	14.8	74.0	811	4	US-09-040-229B-5	Sequence 5, Appl
C 80	14.8	74.0	811	4	US-09-040-229B-5	Sequence 5, Appl
C 81	14.8	74.0	1027	4	US-09-684-405-8	Sequence 8, Appl
C 82	14.8	74.0	1027	4	US-09-684-405-8	Sequence 8, Appl
C 83	14.8	74.0	1781	4	US-09-649-747A-20	Sequence 20, Appl
C 84	14.8	74.0	1781	4	US-09-649-747A-20	Sequence 20, Appl
C 85	14.8	74.0	1949	4	US-09-649-747A-12	Sequence 12, Appl
C 86	14.8	74.0	1949	4	US-09-649-747A-12	Sequence 12, Appl
C 87	14.8	74.0	87350	3	US-08-781-891-79	Sequence 79, Appl
C 88	14.8	74.0	87350	3	US-08-781-891-79	Sequence 79, Appl
C 89	14.8	74.0	87350	4	US-09-618-166-79	Sequence 79, Appl
C 90	14.8	74.0	87350	4	US-09-618-166-79	Sequence 79, Appl
C 91	14.8	74.0	87543	4	US-09-791-211-3	Sequence 3, Appl
C 92	14.8	74.0	87543	4	US-09-791-211-3	Sequence 3, Appl
C 93	14.4	72.0	5066	4	US-09-824-574-1	Sequence 1, Appl
C 94	14.4	72.0	5066	4	US-09-824-574-1	Sequence 1, Appl
C 95	14.2	71.0	303	4	US-09-252-991A-7045	Sequence 7045, Ap
C 96	14.2	71.0	303	4	US-09-252-991A-7045	Sequence 7045, Ap
C 97	14.2	71.0	393	4	US-09-252-991A-4966	Sequence 4966, Ap
C 98	14.2	71.0	393	4	US-09-252-991A-4966	Sequence 4966, Ap
C 99	14.2	71.0	568	4	US-09-513-999C-1079	Sequence 1079, Ap
C 100	14.2	71.0	568	4	US-09-513-999C-1079	Sequence 1079, Ap

101	14.2	71.0	693	4	US-09-252-991A-13194	Sequence 13194, A	c 174	13.8	69.0	578	4	US-09-604-287A-150	Sequence 150, App
c 102	14.2	71.0	693	4	US-09-252-991A-13194	Sequence 13194, A	c 175	13.8	69.0	578	4	US-09-285-480-150	Sequence 150, App
c 103	14.2	71.0	726	4	US-09-252-991A-4997	Sequence 4997, Ap	c 176	13.8	69.0	578	4	US-09-285-480-150	Sequence 150, App
c 104	14.2	71.0	726	4	US-09-252-991A-4997	Sequence 4997, Ap	c 177	13.8	69.0	578	4	US-09-834-759-150	Sequence 150, App
c 105	14.2	71.0	759	4	US-09-252-991A-13714	Sequence 13714, A	c 178	13.8	69.0	578	4	US-09-834-759-150	Sequence 150, App
c 106	14.2	71.0	759	4	US-09-252-991A-13714	Sequence 13714, A	c 179	13.8	69.0	578	4	US-09-590-751A-150	Sequence 150, App
c 107	14.2	71.0	798	4	US-09-252-991A-13429	Sequence 13429, A	c 180	13.8	69.0	578	4	US-09-590-751A-150	Sequence 150, App
c 108	14.2	71.0	798	4	US-09-252-991A-13429	Sequence 13429, A	c 181	13.8	69.0	649	3	US-09-328-111-114	Sequence 114, App
c 109	14.2	71.0	816	4	US-09-252-991A-4951	Sequence 4951, Ap	c 182	13.8	69.0	799	3	US-09-328-111-114	Sequence 114, App
c 110	14.2	71.0	816	4	US-09-252-991A-4951	Sequence 4951, Ap	c 183	13.8	69.0	799	3	US-08-998-416-420	Sequence 420, App
c 111	14.2	71.0	816	4	US-09-252-991A-13914	Sequence 13914, A	c 184	13.8	69.0	799	3	US-08-998-416-420	Sequence 420, App
c 112	14.2	71.0	816	4	US-09-252-991A-13914	Sequence 13914, A	c 185	13.8	69.0	1001	4	US-09-671-317-347	Sequence 347, App
c 113	14.2	71.0	966	4	US-09-252-991A-13437	Sequence 13437, A	c 186	13.8	69.0	1001	4	US-09-671-317-347	Sequence 347, App
c 114	14.2	71.0	966	4	US-09-252-991A-13437	Sequence 13437, A	c 187	13.8	69.0	1656	3	US-09-026-958-1	Sequence 1, Appli
c 115	14.2	71.0	1125	4	US-09-252-991A-11202	Sequence 11202, A	c 188	13.8	69.0	1656	3	US-09-026-958-1	Sequence 1, Appli
c 116	14.2	71.0	1125	4	US-09-252-991A-11202	Sequence 11202, A	c 189	13.8	69.0	1810	4	US-09-800-729-73	Sequence 73, Appl
c 117	14.2	71.0	1134	4	US-09-252-991A-12761	Sequence 12761, A	c 190	13.8	69.0	1810	4	US-09-800-729-73	Sequence 73, Appl
c 118	14.2	71.0	1134	4	US-09-252-991A-12761	Sequence 12761, A	c 191	13.8	69.0	1811	4	US-09-800-729-77	Sequence 77, Appl
c 119	14.2	71.0	1140	1	US-08-314-172A-13	Sequence 13, Appl	c 192	13.8	69.0	1811	4	US-09-800-729-77	Sequence 77, Appl
c 120	14.2	71.0	1140	1	US-08-314-172A-13	Sequence 13, Appl	c 193	13.8	69.0	1964	3	US-09-434-613-2	Sequence 2, Appli
c 121	14.2	71.0	1254	4	US-09-252-991A-10642	Sequence 10642, A	c 194	13.8	69.0	1964	3	US-09-434-613-2	Sequence 2, Appli
c 122	14.2	71.0	1254	4	US-09-252-991A-10642	Sequence 10642, A	c 195	13.8	69.0	1964	4	US-09-963-908-2	Sequence 2, Appli
c 123	14.2	71.0	1308	4	US-09-252-991A-12614	Sequence 12614, A	c 196	13.8	69.0	1964	4	US-09-963-908-2	Sequence 2, Appli
c 124	14.2	71.0	1308	4	US-09-252-991A-12614	Sequence 12614, A	c 197	13.8	69.0	1985	4	US-09-620-312D-812	Sequence 812, App
c 125	14.2	71.0	1488	4	US-09-252-991A-4938	Sequence 4938, Ap	c 198	13.8	69.0	1985	4	US-09-620-312D-812	Sequence 812, App
c 126	14.2	71.0	1488	4	US-09-252-991A-4938	Sequence 4938, Ap	c 199	13.8	69.0	4768	4	US-09-526-193A-16	Sequence 16, Appl
c 127	14.2	71.0	1668	4	US-09-252-991A-13705	Sequence 13705, A	c 200	13.8	69.0	4768	4	US-09-526-193A-16	Sequence 16, Appl
c 128	14.2	71.0	1668	4	US-09-252-991A-13705	Sequence 13705, A	c 201	13.6	68.0	20	3	US-09-210-748A-6	Sequence 6, Appli
c 129	14.2	71.0	1803	3	US-09-369-364A-20	Sequence 20, Appl	c 202	13.6	68.0	20	3	US-09-210-748A-6	Sequence 6, Appli
c 130	14.2	71.0	1803	3	US-09-369-364A-20	Sequence 20, Appl	c 203	13.6	68.0	20	4	US-09-939-581A-6	Sequence 6, Appli
c 131	14.2	71.0	1947	4	US-09-252-991A-7111	Sequence 7111, Ap	c 204	13.6	68.0	132	4	US-09-513-999C-18033	Sequence 18033, A
c 132	14.2	71.0	1947	4	US-09-252-991A-7111	Sequence 7111, Ap	c 205	13.6	68.0	132	4	US-09-513-999C-18033	Sequence 18033, A
c 133	14.2	71.0	2001	4	US-09-213-888-26	Sequence 26, Appl	c 206	13.6	68.0	293	4	US-09-313-294A-5620	Sequence 5620, Ap
c 134	14.2	71.0	2001	4	US-09-213-888-26	Sequence 26, Appl	c 207	13.6	68.0	293	4	US-09-313-294A-5620	Sequence 5620, Ap
c 135	14.2	71.0	2001	4	US-09-328-877D-26	Sequence 26, Appl	c 208	13.6	68.0	341	4	US-09-621-976-18936	Sequence 18936, A
c 136	14.2	71.0	2001	4	US-09-328-877D-26	Sequence 26, Appl	c 209	13.6	68.0	341	4	US-09-621-976-18936	Sequence 18936, A
c 137	14.2	71.0	2010	4	US-09-213-888-24	Sequence 24, Appl	c 210	13.6	68.0	366	4	US-09-513-999C-13574	Sequence 13574, A
c 138	14.2	71.0	2010	4	US-09-213-888-24	Sequence 24, Appl	c 211	13.6	68.0	366	4	US-09-513-999C-13574	Sequence 13574, A
c 139	14.2	71.0	2010	4	US-09-328-877D-24	Sequence 24, Appl	c 212	13.6	68.0	368	4	US-09-404-879A-269	Sequence 269, App
c 140	14.2	71.0	2010	4	US-09-328-877D-24	Sequence 24, Appl	c 213	13.6	68.0	368	4	US-09-404-879A-269	Sequence 269, App
c 141	14.2	71.0	2169	4	US-09-252-991A-10432	Sequence 10432, A	c 214	13.6	68.0	368	4	US-09-404-879A-269	Sequence 269, App
c 142	14.2	71.0	2169	4	US-09-252-991A-10432	Sequence 10432, A	c 215	13.6	68.0	368	4	US-09-338-933-269	Sequence 269, App
c 143	14.2	71.0	2979	4	US-09-252-991A-10524	Sequence 10524, A	c 216	13.6	68.0	368	4	US-09-338-933-269	Sequence 269, App
c 144	14.2	71.0	2979	4	US-09-252-991A-10524	Sequence 10524, A	c 217	13.6	68.0	368	4	US-09-215-681-269	Sequence 269, App
c 145	14.2	71.0	3550	4	US-09-213-888-1	Sequence 1, Appli	c 218	13.6	68.0	368	4	US-09-215-681-269	Sequence 269, App
c 146	14.2	71.0	3550	4	US-09-213-888-1	Sequence 1, Appli	c 219	13.6	68.0	368	4	US-09-215-681-269	Sequence 269, App
c 147	14.2	71.0	3550	4	US-09-328-877D-1	Sequence 1, Appli	c 220	13.6	68.0	368	4	US-09-216-003A-269	Sequence 269, App
c 148	14.2	71.0	3550	4	US-09-328-877D-1	Sequence 1, Appli	c 221	13.6	68.0	368	4	US-09-216-003A-269	Sequence 269, App
c 149	14.2	71.0	3559	4	US-09-800-729-53	Sequence 53, Appl	c 222	13.6	68.0	368	4	US-09-667-857-269	Sequence 269, App
c 150	14.2	71.0	3559	4	US-09-800-729-53	Sequence 53, Appl	c 223	13.6	68.0	429	4	US-09-107-532A-160	Sequence 160, App
c 151	14.2	71.0	5720	4	US-09-800-729-18	Sequence 18, Appl	c 224	13.6	68.0	429	4	US-09-107-532A-160	Sequence 160, App
c 152	14.2	71.0	5720	4	US-09-800-729-18	Sequence 18, Appl	c 225	13.6	68.0	447	4	US-09-621-976-9361	Sequence 9361, Ap
c 153	14.2	71.0	193303	4	US-09-497-855A-37	Sequence 37, Appl	c 226	13.6	68.0	541	4	US-09-621-976-9361	Sequence 9361, Ap
c 154	14.2	71.0	193303	4	US-09-497-855A-37	Sequence 37, Appl	c 227	13.6	68.0	541	4	US-09-621-976-9361	Sequence 9361, Ap
c 155	14.2	71.0	193303	4	US-09-497-855A-44	Sequence 44, Appl	c 228	13.6	68.0	798	4	US-09-621-976-18562	Sequence 18562, A
c 156	14.2	71.0	193303	4	US-09-497-855A-44	Sequence 44, Appl	c 229	13.6	68.0	798	4	US-09-621-976-18562	Sequence 18562, A
c 157	14.2	71.0	536165	4	US-09-214-808-1	Sequence 1, Appli	c 230	13.6	68.0	798	4	US-09-270-767-12368	Sequence 12368, A
c 158	14.2	71.0	536165	4	US-09-214-808-1	Sequence 1, Appli	c 231	13.6	68.0	801	4	US-09-270-767-12368	Sequence 12368, A
c 159	14	70.0	35881	4	US-08-311-731A-127	Sequence 127, App	c 232	13.6	68.0	801	4	US-09-270-767-12368	Sequence 12368, A
c 160	14	70.0	35881	4	US-08-311-731A-127	Sequence 127, App	c 233	13.6	68.0	801	4	US-09-270-767-12368	Sequence 12368, A
c 161	13.8	69.0	435	4	US-08-311-731A-127	Sequence 127, App	c 234	13.6	68.0	834	4	US-09-252-991A-348	Sequence 348, App
c 162	13.8	69.0	435	4	US-08-311-731A-127	Sequence 127, App	c 235	13.6	68.0	834	4	US-09-252-991A-348	Sequence 348, App
c 163	13.8	69.0	578	3	US-09-222-575-150	Sequence 150, App	c 236	13.6	68.0	834	4	US-09-252-991A-348	Sequence 348, App
c 164	13.8	69.0	578	3	US-09-222-575-150	Sequence 150, App	c 237	13.6	68.0	901	3	US-09-292-858B-5	Sequence 5, Appli
c 165	13.8	69.0	578	4	US-09-389-661-150	Sequence 150, App	c 238	13.6	68.0	901	3	US-09-292-858B-5	Sequence 5, Appli
c 166	13.8	69.0	578	4	US-09-389-661-150	Sequence 150, App	c 239	13.6	68.0	901	3	US-09-328-111-2	Sequence 2, Appli
c 167	13.8	69.0	578	4	US-09-620-405B-150	Sequence 150, App	c 240	13.6	68.0	945	4	US-09-107-532A-3562	Sequence 3562, Ap
c 168	13.8	69.0	578	4	US-09-620-405B-150	Sequence 150, App	c 241	13.6	68.0	945	4	US-09-107-532A-3562	Sequence 3562, Ap
c 169	13.8	69.0	578	4	US-09-339-338-150	Sequence 150, App	c 242	13.6	68.0	969	4	US-09-252-991A-4308	Sequence 4308, Ap
c 170	13.8	69.0	578	4	US-09-339-338-150	Sequence 150, App	c 243	13.6	68.0	969	4	US-09-252-991A-4308	Sequence 4308, Ap
c 171	13.8	69.0	578	4	US-09-433-826B-150	Sequence 150, App	c 244	13.6	68.0	999	4	US-09-724-797-17	Sequence 17, Appl
c 172	13.8	69.0	578	4	US-09-433-826B-150	Sequence 150, App	c 245	13.6	68.0	1047	4	US-09-724-797-17	Sequence 17, Appl
c 173	13.8	69.0	578	4	US-09-604-287A-150	Sequence 150, App	c 246	13.6	68.0	1047	4	US-09-604-287A-150	Sequence 150, App

247	13.6	68.0	1080	4	US-09-891-641-63	Sequence 63, Appl	C 320	13.6	68.0	8878	1	US-08-759-444-2	Sequence 2, Appl
C 248	13.6	68.0	1080	4	US-09-891-641-63	Sequence 63, Appl	C 321	13.6	68.0	9880	3	US-08-680-897-1	Sequence 1, Appl
C 249	13.6	68.0	1200	4	US-09-620-312D-1095	Sequence 1095, Ap	C 322	13.6	68.0	9880	3	US-08-680-897-1	Sequence 1, Appl
C 250	13.6	68.0	1200	4	US-09-620-312D-1095	Sequence 1095, Ap	C 323	13.6	68.0	11187	4	US-09-422-936-61	Sequence 61, Appl
C 251	13.6	68.0	1392	4	US-09-252-991A-12249	Sequence 12249, A	C 324	13.6	68.0	11187	4	US-09-422-936-61	Sequence 61, Appl
C 252	13.6	68.0	1392	4	US-09-252-991A-12249	Sequence 12249, A	C 325	13.6	68.0	11863	4	US-09-814-915A-83	Sequence 83, Appl
C 253	13.6	68.0	1401	4	US-09-252-991A-320	Sequence 320, App	C 326	13.6	68.0	11863	4	US-09-814-915A-83	Sequence 83, Appl
C 254	13.6	68.0	1401	4	US-09-252-991A-320	Sequence 320, App	C 327	13.6	68.0	15602	4	US-09-844-634-17	Sequence 17, Appl
C 255	13.6	68.0	1407	4	US-09-252-991A-12361	Sequence 12361, A	C 328	13.6	68.0	15602	4	US-09-844-634-17	Sequence 17, Appl
C 256	13.6	68.0	1407	4	US-09-252-991A-12361	Sequence 12361, A	C 329	13.6	68.0	37950	3	US-09-338-907-183	Sequence 183, App
C 257	13.6	68.0	1530	4	US-09-252-991A-12109	Sequence 12109, A	C 330	13.6	68.0	37950	3	US-09-338-907-183	Sequence 183, App
C 258	13.6	68.0	1530	4	US-09-252-991A-12109	Sequence 12109, A	C 331	13.6	68.0	37950	3	US-09-218-207-183	Sequence 183, App
C 259	13.6	68.0	1596	4	US-09-252-991A-9478	Sequence 9478, Ap	C 332	13.6	68.0	37950	3	US-09-218-207-183	Sequence 183, App
C 260	13.6	68.0	1596	4	US-09-252-991A-9478	Sequence 9478, Ap	C 333	13.6	68.0	55298	4	US-09-491-356C-1	Sequence 1, Appl
C 261	13.6	68.0	1620	3	US-09-360-197-3	Sequence 3, Appl	C 334	13.6	68.0	55298	4	US-09-491-356C-1	Sequence 1, Appl
C 262	13.6	68.0	1620	3	US-09-360-197-3	Sequence 3, Appl	C 335	13.4	67.0	30	3	US-08-648-506-1	Sequence 1, Appl
C 263	13.6	68.0	1650	4	US-09-252-991A-9352	Sequence 9352, Ap	C 336	13.4	67.0	30	3	US-08-648-506-1	Sequence 1, Appl
C 264	13.6	68.0	1650	4	US-09-252-991A-9352	Sequence 9352, Ap	C 337	13.4	67.0	30	3	US-09-277-355-1	Sequence 1, Appl
C 265	13.6	68.0	1778	4	US-09-799-451-790	Sequence 790, App	C 338	13.4	67.0	30	3	US-09-277-355-1	Sequence 1, Appl
C 266	13.6	68.0	1778	4	US-09-799-451-790	Sequence 790, App	C 339	13.4	67.0	366	4	US-09-569-852B-4	Sequence 4, Appl
C 267	13.6	68.0	2089	1	US-08-552-142A-1	Sequence 1, Appl	C 340	13.4	67.0	366	4	US-09-569-852B-4	Sequence 4, Appl
C 268	13.6	68.0	2089	1	US-08-552-142A-1	Sequence 1, Appl	C 341	13.4	67.0	417	4	US-09-252-991A-3455	Sequence 3455, Ap
C 269	13.6	68.0	2089	1	US-08-910-973-1	Sequence 1, Appl	C 342	13.4	67.0	417	4	US-09-252-991A-3455	Sequence 3455, Ap
C 270	13.6	68.0	2089	1	US-08-910-973-1	Sequence 1, Appl	C 343	13.4	67.0	543	4	US-09-328-352-9957	Sequence 3957, Ap
C 271	13.6	68.0	2089	4	US-09-499-227-1	Sequence 1, Appl	C 344	13.4	67.0	543	4	US-09-328-352-9957	Sequence 3957, Ap
C 272	13.6	68.0	2089	4	US-09-499-227-1	Sequence 1, Appl	C 345	13.4	67.0	603	3	US-08-998-416-1152	Sequence 1152, Ap
C 273	13.6	68.0	2089	5	PCT-US95-05741-1	Sequence 1, Appl	C 346	13.4	67.0	603	3	US-08-998-416-1152	Sequence 1152, Ap
C 274	13.6	68.0	2089	5	PCT-US95-05741-1	Sequence 1, Appl	C 347	13.4	67.0	1026	4	US-09-252-991A-3491	Sequence 3491, Ap
C 275	13.6	68.0	2091	4	US-09-252-991A-12312	Sequence 12312, A	C 348	13.4	67.0	1026	4	US-09-252-991A-3491	Sequence 3491, Ap
C 276	13.6	68.0	2091	4	US-09-252-991A-12312	Sequence 12312, A	C 349	13.4	67.0	1027	4	US-09-674-741-9	Sequence 9, Appl
C 277	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 350	13.4	67.0	1027	4	US-09-674-741-9	Sequence 9, Appl
C 278	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 351	13.4	67.0	1142	3	US-09-227-357-14	Sequence 14, Appl
C 279	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 352	13.4	67.0	1142	3	US-09-227-357-14	Sequence 14, Appl
C 280	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 353	13.4	67.0	1172	4	US-09-311-021-149	Sequence 149, App
C 281	13.6	68.0	2224	4	US-09-800-909-1	Sequence 1, Appl	C 354	13.4	67.0	1172	4	US-09-311-021-149	Sequence 149, App
C 282	13.6	68.0	2224	4	US-09-800-909-1	Sequence 1, Appl	C 355	13.4	67.0	1302	4	US-09-252-991A-3473	Sequence 3473, Ap
C 283	13.6	68.0	2224	4	US-09-800-909-1	Sequence 1, Appl	C 356	13.4	67.0	1302	4	US-09-252-991A-3473	Sequence 3473, Ap
C 284	13.6	68.0	2224	4	US-09-800-908-2	Sequence 2, Appl	C 357	13.4	67.0	1332	4	US-09-252-991A-3464	Sequence 3464, Ap
C 285	13.6	68.0	2266	2	US-08-724-394A-18	Sequence 18, Appl	C 358	13.4	67.0	1332	4	US-09-252-991A-3464	Sequence 3464, Ap
C 286	13.6	68.0	2266	2	US-08-724-394A-18	Sequence 18, Appl	C 359	13.4	67.0	1722	4	US-09-252-991A-3485	Sequence 3485, Ap
C 287	13.6	68.0	2431	4	US-09-221-017B-351	Sequence 351, App	C 360	13.4	67.0	1722	4	US-09-252-991A-3485	Sequence 3485, Ap
C 288	13.6	68.0	2431	4	US-09-221-017B-351	Sequence 351, App	C 361	13.4	67.0	1762	4	US-09-799-451-122	Sequence 622, App
C 289	13.6	68.0	2490	4	US-09-252-991A-10579	Sequence 10579, A	C 362	13.4	67.0	1762	4	US-09-799-451-122	Sequence 622, App
C 290	13.6	68.0	2490	4	US-09-252-991A-10579	Sequence 10579, A	C 363	13.4	67.0	1904	4	US-10-140-002-99	Sequence 99, Appl
C 291	13.6	68.0	2613	4	US-09-856-937A-1	Sequence 1, Appl	C 364	13.4	67.0	1904	4	US-10-140-002-99	Sequence 99, Appl
C 292	13.6	68.0	2613	4	US-09-856-937A-1	Sequence 1, Appl	C 365	13.4	67.0	2907	4	US-09-023-655-1053	Sequence 1053, Ap
C 293	13.6	68.0	2818	4	US-09-491-356C-4	Sequence 4, Appl	C 366	13.4	67.0	2907	4	US-09-023-655-1053	Sequence 1053, Ap
C 294	13.6	68.0	2818	4	US-09-491-356C-4	Sequence 4, Appl	C 367	13.4	67.0	4517	3	US-09-140-804-9	Sequence 9, Appl
C 295	13.6	68.0	2974	1	US-08-290-978A-4	Sequence 4, Appl	C 368	13.4	67.0	4517	3	US-09-140-804-9	Sequence 9, Appl
C 296	13.6	68.0	2974	1	US-08-290-978A-4	Sequence 4, Appl	C 369	13.4	67.0	4517	4	US-09-686-838B-9	Sequence 9, Appl
C 297	13.6	68.0	2974	2	US-08-780-869-4	Sequence 4, Appl	C 370	13.4	67.0	4517	4	US-09-686-838B-9	Sequence 9, Appl
C 298	13.6	68.0	2974	2	US-08-780-869-4	Sequence 4, Appl	C 371	13.4	67.0	4517	4	US-09-776-976-5	Sequence 5, Appl
C 299	13.6	68.0	3323	4	US-10-140-002-167	Sequence 167, App	C 372	13.4	67.0	4517	4	US-09-776-976-5	Sequence 5, Appl
C 300	13.6	68.0	3323	4	US-10-140-002-167	Sequence 167, App	C 373	13.4	67.0	4517	4	US-09-909-547-5	Sequence 5, Appl
C 301	13.6	68.0	3384	4	US-09-252-991A-10278	Sequence 10278, A	C 374	13.4	67.0	4517	4	US-09-909-547-5	Sequence 5, Appl
C 302	13.6	68.0	3384	4	US-09-252-991A-10278	Sequence 10278, A	C 375	13.4	67.0	5526	3	US-08-751-359-21	Sequence 21, Appl
C 303	13.6	68.0	3403	4	US-09-023-655-737	Sequence 737, App	C 376	13.4	67.0	5526	3	US-08-751-359-21	Sequence 21, Appl
C 304	13.6	68.0	3403	4	US-09-023-655-737	Sequence 737, App	C 377	13.4	67.0	5526	3	US-08-907-146-21	Sequence 21, Appl
C 305	13.6	68.0	3492	4	US-09-023-655-1219	Sequence 1219, App	C 378	13.4	67.0	5526	3	US-08-907-146-21	Sequence 21, Appl
C 306	13.6	68.0	3492	4	US-09-023-655-1219	Sequence 1219, App	C 379	13.4	67.0	28438	4	US-09-820-790B-3	Sequence 3, Appl
C 307	13.6	68.0	3683	4	US-09-844-634-3	Sequence 3, Appl	C 380	13.4	67.0	28438	4	US-09-820-790B-3	Sequence 3, Appl
C 308	13.6	68.0	3683	4	US-09-844-634-3	Sequence 3, Appl	C 381	13.2	66.0	20	4	US-09-702-327-44	Sequence 44, Appl
C 309	13.6	68.0	3683	4	US-09-844-634-3	Sequence 3, Appl	C 382	13.2	66.0	20	4	US-09-702-327-44	Sequence 44, Appl
C 310	13.6	68.0	3683	4	US-09-968-455-1	Sequence 1, Appl	C 383	13.2	66.0	137	3	US-08-998-416-465	Sequence 465, App
C 311	13.6	68.0	4233	4	US-09-968-455-1	Sequence 1, Appl	C 384	13.2	66.0	137	3	US-08-998-416-465	Sequence 465, App
C 312	13.6	68.0	4233	4	US-09-968-455-1	Sequence 1, Appl	C 385	13.2	66.0	138	3	US-08-998-416-464	Sequence 464, App
C 313	13.6	68.0	4436	4	US-09-491-356C-5	Sequence 5, Appl	C 386	13.2	66.0	138	3	US-08-998-416-464	Sequence 464, App
C 314	13.6	68.0	4436	4	US-09-491-356C-5	Sequence 5, Appl	C 387	13.2	66.0	234	4	US-09-513-999C-14238	Sequence 14238, A
C 315	13.6	68.0	5140	3	US-09-491-356C-6	Sequence 6, Appl	C 388	13.2	66.0	234	4	US-09-513-999C-14238	Sequence 14238, A
C 316	13.6	68.0	5140	3	US-09-333-214-2	Sequence 2, Appl	C 389	13.2	66.0	263	4	US-09-636-215-682	Sequence 682, App
C 317	13.6	68.0	7244	4	US-09-333-214-2	Sequence 2, Appl	C 390	13.2	66.0	263	4	US-09-636-215-682	Sequence 682, App
C 318	13.6	68.0	7244	4	US-09-774-528-143	Sequence 143, App	C 391	13.2	66.0	263	4	US-09-636-215-682	Sequence 682, App
C 319	13.6	68.0	8878	1	US-08-759-444-2	Sequence 2, Appl	C 392	13.2	66.0	263	4	US-09-685-166A-682	Sequence 682, App

393	13.2	66.0	265	4	US-09-679-426-682	Sequence 682, App	1649	4	US-09-484-970B-72	Sequence 72, Appl
394	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App	1649	4	US-09-484-970B-72	Sequence 72, Appl
395	13.2	66.0	282	3	US-09-042-353-94	Sequence 94, Appl	1654	4	US-09-921-017B-326	Sequence 326, Appl
396	13.2	66.0	282	3	US-09-042-353-94	Sequence 94, Appl	1654	4	US-09-921-017B-326	Sequence 326, Appl
397	13.2	66.0	282	3	US-08-758-417A-358	Sequence 358, App	1666	3	US-09-360-197-5	Sequence 5, Appl
398	13.2	66.0	282	3	US-08-758-417A-358	Sequence 358, App	1666	3	US-09-360-197-5	Sequence 5, Appl
399	13.2	66.0	295	4	US-09-313-284A-2859	Sequence 2859, App	1674	4	US-09-620-312D-593	Sequence 593, App
400	13.2	66.0	295	4	US-09-313-284A-2859	Sequence 2859, App	1674	4	US-09-620-312D-593	Sequence 593, App
401	13.2	66.0	295	4	US-09-313-284A-2859	Sequence 2859, App	1674	4	US-09-620-312D-593	Sequence 593, App
402	13.2	66.0	295	4	US-09-313-284A-2859	Sequence 2859, App	1674	4	US-09-620-312D-593	Sequence 593, App
403	13.2	66.0	340	4	US-09-513-999C-34977	Sequence 34977, A	1785	4	US-09-489-039A-5174	Sequence 5174, App
404	13.2	66.0	340	4	US-09-513-999C-34977	Sequence 34977, A	1785	4	US-09-489-039A-5174	Sequence 5174, App
405	13.2	66.0	377	4	US-09-513-999C-1087	Sequence 1087, App	1914	4	US-09-252-991A-1914	Sequence 1914, App
406	13.2	66.0	377	4	US-09-404-879A-259	Sequence 259, App	2176	4	US-09-843-250-8	Sequence 8, Appl
407	13.2	66.0	377	4	US-09-404-879A-259	Sequence 259, App	2176	4	US-09-843-250-8	Sequence 8, Appl
408	13.2	66.0	377	4	US-09-338-933-259	Sequence 259, App	2400	1	US-08-785-052-3	Sequence 3, Appl
409	13.2	66.0	377	4	US-09-338-933-259	Sequence 259, App	2400	1	US-08-785-052-3	Sequence 3, Appl
410	13.2	66.0	377	4	US-09-215-681-259	Sequence 259, App	2400	2	US-08-913-581-3	Sequence 3, Appl
411	13.2	66.0	377	4	US-09-215-681-259	Sequence 259, App	2400	2	US-08-913-581-3	Sequence 3, Appl
412	13.2	66.0	377	4	US-09-216-003A-259	Sequence 259, App	2571	4	US-09-091-885-1	Sequence 1, Appl
413	13.2	66.0	377	4	US-09-216-003A-259	Sequence 259, App	2571	4	US-09-091-885-1	Sequence 1, Appl
414	13.2	66.0	377	4	US-09-667-857-259	Sequence 259, App	4355	4	US-09-843-250-7	Sequence 8190, App
415	13.2	66.0	427	4	US-09-667-857-259	Sequence 259, App	4355	4	US-09-843-250-7	Sequence 8190, App
416	13.2	66.0	427	4	US-09-667-857-259	Sequence 259, App	4355	4	US-09-843-250-7	Sequence 8190, App
417	13.2	66.0	427	4	US-09-667-857-259	Sequence 259, App	4355	4	US-09-843-250-7	Sequence 8190, App
418	13.2	66.0	446	4	US-09-513-999C-2207	Sequence 2207, App	4848	3	US-08-955-957A-1	Sequence 7, Appl
419	13.2	66.0	446	4	US-09-513-999C-2207	Sequence 2207, App	4848	3	US-08-955-957A-1	Sequence 7, Appl
420	13.2	66.0	450	4	US-09-621-976-1484	Sequence 1484, App	4848	3	US-08-955-957A-1	Sequence 1, Appl
421	13.2	66.0	450	4	US-09-621-976-1484	Sequence 1484, App	4848	3	US-08-955-957A-1	Sequence 1, Appl
422	13.2	66.0	452	4	US-09-513-999C-4050	Sequence 4050, App	4848	3	US-08-955-957A-1	Sequence 4, Appl
423	13.2	66.0	452	4	US-09-513-999C-4050	Sequence 4050, App	4848	3	US-08-955-957A-1	Sequence 4, Appl
424	13.2	66.0	765	3	US-08-998-416-480	Sequence 480, App	4848	3	US-08-955-957A-1	Sequence 6, Appl
425	13.2	66.0	765</							



INFORMATION FOR SEQ ID NO: 162:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Primer"  
US-08-904-901-162

Query Match 100.0%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 1 GGACATGCCCGGCATGTCC 20

## RESULT 2

US-08-904-901-162/c  
Sequence 162, Application US/08904901  
Patent No. 5998383  
GENERAL INFORMATION:  
APPLICANT: Wright, Jim A.  
APPLICANT: Young, Aiping H.  
TITLE OF INVENTION: ANTITUMOR ANTISENSE SEQUENCES DIRECTED  
TITLE OF INVENTION: AGAINST RIBONUCLEOTIDE REDUCTASE  
NUMBER OF SEQUENCES: 163  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KOHN & ASSOCIATES  
STREET: 30500 No. 5998383thwestern Hwy. Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: US  
ZIP: 48334  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/904,901  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohn, Kenneth I.  
REGISTRATION NUMBER: 30,955  
REFERENCE/DOCKET NUMBER: 0227,00004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (248) 539-5050  
TELEFAX: (248) 539-5055  
INFORMATION FOR SEQ ID NO: 162:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Primer"  
US-08-904-901-162

Query Match 100.0%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 20 GGACATGCCCGGCATGTCC 1

RESULT 3  
US-09-249-730-162

Sequence 162, Application US/09249730  
Patent No. 6121000  
GENERAL INFORMATION:  
APPLICANT: WRIGHT, Jim A.  
APPLICANT: YOUNG, Aiping H.  
TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and  
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase  
FILE REFERENCE: 032396-040  
CURRENT APPLICATION NUMBER: US/09/249,730  
CURRENT FILING DATE: 1999-02-11  
NUMBER OF SEQ ID NOS: 220  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO: 162  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Human  
US-09-249-730-162

Query Match 100.0%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 1 GGACATGCCCGGCATGTCC 20

## RESULT 4

US-09-249-730-162/c  
Sequence 162, Application US/09249730  
Patent No. 6121000  
GENERAL INFORMATION:  
APPLICANT: WRIGHT, Jim A.  
APPLICANT: YOUNG, Aiping H.  
TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and  
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase  
FILE REFERENCE: 032396-040  
CURRENT APPLICATION NUMBER: US/09/249,730  
CURRENT FILING DATE: 1999-02-11  
NUMBER OF SEQ ID NOS: 220  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO: 162  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Human  
US-09-249-730-162

Query Match 100.0%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 20 GGACATGCCCGGCATGTCC 1

## RESULT 5

US-09-249-247-162  
Sequence 162, Application US/09249247  
Patent No. 6593305  
GENERAL INFORMATION:  
APPLICANT: WRIGHT, Jim A.  
APPLICANT: YOUNG, Aiping H.  
TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and  
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase  
FILE REFERENCE: 032396-023  
CURRENT APPLICATION NUMBER: US/09/249,247  
CURRENT FILING DATE: 1999-02-11  
EARLIER APPLICATION NUMBER: US 60/023,040  
EARLIER FILING DATE: 1996-08-02  
EARLIER APPLICATION NUMBER: US 60/039,959  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: US 08/904,901

EARLIER FILING DATE: 1997-08-01  
NUMBER OF SEQ ID NOS: 220  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 162  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Human  
US-09-249-247-162

Query Match  
Best Local Similarity 100.0%; Score 20; DB 4; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 1 GGACATGCCCGGCATGTCC 20

RESULT 6  
US-09-249-247-162/c  
Sequence 162, Application US/09249247  
Patent No. 6593305  
GENERAL INFORMATION:

APPLICANT: WRIGHT, Jim A.  
APPLICANT: YOUNG, Aiping H.  
TITLE OF INVENTION: Anticancer Antisense Sequences Directed Against R1 and  
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase  
FILE REFERENCE: 032396-023  
CURRENT FILING DATE: 1999-02-11  
EARLIER FILING DATE: 1999-02-11  
EARLIER FILING DATE: 1996-08-02  
EARLIER FILING DATE: 1996-08-02  
EARLIER FILING DATE: 1997-03-07  
EARLIER FILING DATE: 1997-03-07  
EARLIER FILING DATE: 1997-08-01  
NUMBER OF SEQ ID NOS: 220  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 162  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Human  
US-09-249-247-162

Query Match  
Best Local Similarity 100.0%; Score 20; DB 4; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 20 GGACATGCCCGGCATGTCC 1

RESULT 7  
US-08-767-942A-40  
Sequence 40, Application US/08767942A  
Patent No. 6068982  
GENERAL INFORMATION:

APPLICANT: Rolfe, Mark  
APPLICANT: Chiu, M. Isabel  
APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UBIQUITIN CONUGATING ENZYMES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 17-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-767-942A-40

Query Match  
Best Local Similarity 100.0%; Score 20; DB 3; Length 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 6 GGACATGCCCGGCATGTCC 25

RESULT 8  
US-08-767-942A-40/c  
Sequence 40, Application US/08767942A  
Patent No. 6068982  
GENERAL INFORMATION:

APPLICANT: Rolfe, Mark  
APPLICANT: Chiu, M. Isabel  
APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UBIQUITIN CONUGATING ENZYMES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 17-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-767-942A-40

Query Match 100.0%; Score 20; DB 3; Length 26;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
DB 25 GGACATGCCCGGGCATGTCC 6

## RESULT 9

US-08-767-942A-41  
Sequence 41, Application US/08767942A  
Patent No. 6068982  
GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Chiu, M. Isabel  
APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UBIOUITIN CONJUGATING ENZYMES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FOLEY, HONG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,942A  
FILING DATE: 17-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.04  
TELEPHONE: 617-832-1000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-767-942A-41

Query Match 100.0%; Score 20; DB 3; Length 27;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
DB 7 GGACATGCCCGGGCATGTCC 26

## RESULT 10

US-08-767-942A-41/C  
Sequence 41, Application US/08767942A  
Patent No. 6068982  
GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Chiu, M. Isabel

APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UBIOUITIN CONJUGATING ENZYMES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FOLEY, HONG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,942A  
FILING DATE: 17-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.04  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-767-942A-41

Query Match 100.0%; Score 20; DB 3; Length 27;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
DB 26 GGACATGCCCGGGCATGTCC 7

## RESULT 11

US-08-347-792-18  
Sequence 18, Application US/08347792  
Patent No. 5573925  
GENERAL INFORMATION:  
APPLICANT: Halazonec, Thanos D.  
TITLE OF INVENTION: p53 Proteins With Altered  
TITLE OF INVENTION: Tetramerization Domains  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/347,792  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST58USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-347-792-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
|||  
Db 3 GGGCATGTCCGGGCATGTCC 22

RESULT 12  
US-08-347-792-18/c  
Sequence 18, Application US/08347792  
Patent No. 5573925  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 Proteins With Altered  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/347,792  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST58USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-347-792-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
|||  
Db 22 GGACATGCCCGGCATGTCC 3

RESULT 13  
US-08-431-357-18  
Sequence 18, Application US/08431357  
Patent No. 5721340  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 Proteins With Altered  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,357  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/347,792  
FILING DATE: 28-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST58USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-431-357-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
|||  
Db 3 GGGCATGTCCGGGCATGTCC 22

RESULT 14  
US-08-431-357-18/c  
Sequence 18, Application US/08431357  
Patent No. 5721340  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 Proteins With Altered  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,357  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/347,792  
FILING DATE: 28-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST58USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-431-357-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGACATGTC 20  
DB 22 GGACATGCCCGGACATGCC 3

RESULT 15  
US-08-697-221-29  
Sequence 29, Application US/08697221  
Patent No. 5847083  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: Modified p53 Constructs and Uses  
NUMBER OF INVENTION: Therefor  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,221  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,802  
FILING DATE: 22-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kodroff, Cathy A.  
REGISTRATION NUMBER: 33,980  
REFERENCE/DOCKET NUMBER: WST64AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double

TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
US-08-697-221-29

Query Match 84.0%; Score 16.8; DB 2; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGACATGTC 20  
DB 3 GGACATGCCCGGACATGTC 22

RESULT 16  
US-08-697-221-29/c  
Sequence 29, Application US/08697221  
Patent No. 5847083  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: Modified p53 Constructs and Uses  
NUMBER OF INVENTION: Therefor  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,221  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,802  
FILING DATE: 22-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kodroff, Cathy A.  
REGISTRATION NUMBER: 33,980  
REFERENCE/DOCKET NUMBER: WST64AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
US-08-697-221-29

Query Match 84.0%; Score 16.8; DB 2; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGACATGTC 20  
DB 22 GGACATGCCCGGACATGCC 3

RESULT 17  
US-08-392-542-31  
Sequence 31, Application US/08392542  
Patent No. 6169073  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos  
ATTORNEY/AGENT INFORMATION:  
NAME: Hartwig, Wolfgang

```

; TITLE OF INVENTION: Peptides nad Peptidomimetics with
; TITLE OF INVENTION: Structural Similarity to Human p53 That Activate p53
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,542
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9100
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-392-542-31
;
; Query Match      84.0%; Score 16.8; DB 3; Length 30;
; Best Local Similarity 90.0%; Pred. No. 52;
; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 1 GGACATGCCCGGCGCATGTCC 20
; Db 3 GGCGATGTCCGGCGCATGTCC 22
;
; RESULT 18
; US-08-392-542-31/c
; Sequence 31, Application US/08392542
; Patent No. 6169073
; GENERAL INFORMATION:
; APPLICANT: Halazonetis, Thanos
; APPLICANT: Hartwig, Wolfgang
; TITLE OF INVENTION: Peptides nad Peptidomimetics with
; TITLE OF INVENTION: Structural Similarity to Human p53 That Activate p53
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,542
; FILING DATE:

```

```

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 0486,48439
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9100
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-392-542-31
;
; Query Match      84.0%; Score 16.8; DB 3; Length 30;
; Best Local Similarity 90.0%; Pred. No. 52;
; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 1 GGACATGCCCGGCGCATGTCC 20
; Db 22 GGACATGCCCGGCGCATGCC 3
;
; RESULT 19
; US-08-894-327-31
; Sequence 31, Application US/08894327
; Patent No. 6245886
; GENERAL INFORMATION:
; APPLICANT: Halazonetis, Thanos
; APPLICANT: Hartwig, Wolfgang
; TITLE OF INVENTION: Peptides and peptidomimetics with
; TITLE OF INVENTION: structural similarity to human p53 that activate p53
; FILE REFERENCE: 2973,19998
; CURRENT APPLICATION NUMBER: US/08/894,327
; CURRENT FILING DATE: 1997-12-04
; EARLIER APPLICATION NUMBER: pctus96/01535
; EARLIER FILING DATE: 1996-02-16
; EARLIER APPLICATION NUMBER: 08/392,542
; EARLIER FILING DATE: 1995-02-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
;
; US-08-894-327-31
;
; Query Match      84.0%; Score 16.8; DB 3; Length 30;
; Best Local Similarity 90.0%; Pred. No. 52;
; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 1 GGACATGCCCGGCGCATGTCC 20
; Db 3 GGCGATGTCCGGCGCATGTCC 22
;
; RESULT 20
; US-08-894-327-31/c
; Sequence 31, Application US/08894327
; Patent No. 6245886
; GENERAL INFORMATION:
; APPLICANT: Halazonetis, Thanos
; APPLICANT: Hartwig, Wolfgang
; TITLE OF INVENTION: Peptides and peptidomimetics with
; TITLE OF INVENTION: structural similarity to human p53 that activate p53
; FILE REFERENCE: 2973,19998

```

CURRENT APPLICATION NUMBER: US/08/894,327  
CURRENT FILING DATE: 1997-12-04  
EARLIER APPLICATION NUMBER: pctus96/01535  
EARLIER FILING DATE: 1996-02-16  
EARLIER APPLICATION NUMBER: 08/392,542  
EARLIER FILING DATE: 1995-02-16  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 31  
LENGTH: 30  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-894-327-31

Query Match 84.0%; Score 16.8; DB 3; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCGCGGATGTC 20  
22 GGACATGCGCGGATGTC 3

RESULT 21  
US-09-685-027-31  
Sequence 31, Application US/09685027  
Patent No. 6420118  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos  
Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
Structural Similarity to Human p53 That Activate p53  
Function  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: District of Columbia  
COUNTRY: U.S.  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/685,027  
FILING DATE: 10-Oct-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/392,542  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potosrke, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 0486,48439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-09-685-027-31  
Query Match 84.0%; Score 16.8; DB 4; Length 30;

Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCGCGGATGTC 20  
Db 3 GGACATGTCGGGATGTC 22

RESULT 22  
US-09-685-027-31/C  
Sequence 31, Application US/09685027  
Patent No. 6420118  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos  
Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
Structural Similarity to Human p53 That Activate p53  
Function  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: District of Columbia  
COUNTRY: U.S.  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/685,027  
FILING DATE: 10-Oct-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/392,542  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potosrke, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 0486,48439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-09-685-027-31  
Query Match 84.0%; Score 16.8; DB 4; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCGCGGATGTC 20  
Db 22 GGACATGCGCGGATGTC 3

RESULT 23  
US-09-829-922-31  
Sequence 31, Application US/09829922  
Patent No. 6784157  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos  
Hartwig, Wolfgang

```

; TITLE OF INVENTION: Peptides and peptidomimetics with
; TITLE OF INVENTION: structural similarity to human p53 that activate p53
; FILE REFERENCE: 2973.19998
; CURRENT APPLICATION NUMBER: US/09/829,922
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 08/894,327
; PRIOR FILING DATE: 1997-12-04
; PRIOR APPLICATION NUMBER: patus96/01535
; PRIOR FILING DATE: 1996-02-16
; PRIOR APPLICATION NUMBER: 08/392,542
; PRIOR FILING DATE: 1995-02-16
; SOFTWARE: Patent Release #1.0, Version #1.30
; SEQ ID NO 31
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-829-922-31
```

```

Query Match      84.0%; Score 16.8; DB 4; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      1 GGACATGCCCGGGCATGTCC 20
        ||| ||| ||| ||| ||| |||
DB      3 GGGCATGTCCGGGCATGTCC 22
```

```

RESULT 24
US-09-829-922-31/c
; Sequence 31, Application US/09829922
; Patent No. 6784157
; GENERAL INFORMATION:
; APPLICANT: Halazometis, Thanos
; TITLE OF INVENTION: peptides and peptidomimetics with
; TITLE OF INVENTION: structural similarity to human p53 that activate p53
; FILE REFERENCE: 2973.19998
; CURRENT APPLICATION NUMBER: US/09/829,922
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 08/894,327
; PRIOR FILING DATE: 1997-12-04
; PRIOR APPLICATION NUMBER: patus96/01535
; PRIOR FILING DATE: 1996-02-16
; PRIOR APPLICATION NUMBER: 08/392,542
; PRIOR FILING DATE: 1995-02-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-829-922-31
```

```

Query Match      84.0%; Score 16.8; DB 4; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      1 GGACATGCCCGGGCATGTCC 20
        ||| ||| ||| ||| ||| |||
DB      22 GGACATGCCCGGCATGTCC 3
```

```

RESULT 25
PCT-US95-15353-18
; Sequence 18, Application PC/TUS9515353
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of Anatomy
; APPLICANT: and Biology
; APPLICANT: Halazometis, Thanos D.
; TITLE OF INVENTION: p53 Proteins With Altered
```

```

; TITLE OF INVENTION: Tetramerization Domains
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Howson and Howson
; STREET: Spring House Corporate Cntr., PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15353
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/347,792
; FILING DATE: 28-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/431,357
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/456,623
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: MST58CPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US95-15353-18
```

```

Query Match      84.0%; Score 16.8; DB 5; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      1 GGACATGCCCGGGCATGTCC 20
        ||| ||| ||| ||| ||| |||
DB      3 GGGCATGTCCGGGCATGTCC 22
```

```

RESULT 26
PCT-US95-15353-18/c
; Sequence 18, Application PC/TUS9515353
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of Anatomy
; APPLICANT: and Biology
; APPLICANT: Halazometis, Thanos D.
; TITLE OF INVENTION: p53 proteins with altered
; TITLE OF INVENTION: Tetramerization Domains
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Howson and Howson
; STREET: Spring House Corporate Cntr., PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```



SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15353  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/347,792  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/431,357  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/456,623  
FILING DATE: 01-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST58CPCPT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US95-15353-18

Query Match 84.0%; Score 16.8; DB 5; Length 30;  
Best Local Similarity 90.0%; Pred. No. 52;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
Db 22 GGACATGCCCGGCGCATGTCC 3

RESULT 27  
US-09-562-702A-11  
Sequence 11, Application US/09562702A  
Patent No. 6632790  
GENERAL INFORMATION:  
APPLICANT: Yurchenco, Peter  
TITLE OF INVENTION: Laminin 2 and Methods for Its Use  
FILE REFERENCE: 99-274-B  
CURRENT APPLICATION NUMBER: US/09/562,702A  
CURRENT FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/155,945  
PRIOR FILING DATE: 1999-09-24  
PRIOR APPLICATION NUMBER: 60/143,289  
PRIOR FILING DATE: 1999-07-12  
PRIOR APPLICATION NUMBER: 60/139,198  
PRIOR FILING DATE: 1999-06-15  
PRIOR APPLICATION NUMBER: 60/131,720  
PRIOR FILING DATE: 1999-04-30  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 11  
LENGTH: 9391  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(9252)  
US-09-562-702A-11

Query Match 84.0%; Score 16.8; DB 4; Length 9391;  
Best Local Similarity 90.0%; Pred. No. 53;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGACATGCCCGGCGCATGTCC 20

Db 793 GGTCATGCCCGGCGCTTGTC 812

RESULT 28  
US-09-562-702A-11/c  
Sequence 11, Application US/09562702A  
Patent No. 6632790  
GENERAL INFORMATION:  
APPLICANT: Yurchenco, Peter  
TITLE OF INVENTION: Laminin 2 and Methods for Its Use  
FILE REFERENCE: 99-274-B  
CURRENT APPLICATION NUMBER: US/09/562,702A  
CURRENT FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/155,945  
PRIOR FILING DATE: 1999-09-24  
PRIOR APPLICATION NUMBER: 60/143,289  
PRIOR FILING DATE: 1999-07-12  
PRIOR APPLICATION NUMBER: 60/139,198  
PRIOR FILING DATE: 1999-06-15  
PRIOR APPLICATION NUMBER: 60/131,720  
PRIOR FILING DATE: 1999-04-30  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 11  
LENGTH: 9391  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(9252)  
US-09-562-702A-11

Query Match 84.0%; Score 16.8; DB 4; Length 9391;  
Best Local Similarity 90.0%; Pred. No. 53;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
Db 812 GGACATGCCCGGCGCATGTCC 793

RESULT 29  
US-09-562-702A-9  
Sequence 9, Application US/09562702A  
Patent No. 6632790  
GENERAL INFORMATION:  
APPLICANT: Yurchenco, Peter  
TITLE OF INVENTION: Laminin 2 and Methods for Its Use  
FILE REFERENCE: 99-274-B  
CURRENT APPLICATION NUMBER: US/09/562,702A  
CURRENT FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/155,945  
PRIOR FILING DATE: 1999-09-24  
PRIOR APPLICATION NUMBER: 60/143,289  
PRIOR FILING DATE: 1999-07-12  
PRIOR APPLICATION NUMBER: 60/139,198  
PRIOR FILING DATE: 1999-06-15  
PRIOR APPLICATION NUMBER: 60/131,720  
PRIOR FILING DATE: 1999-04-30  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 9  
LENGTH: 9511  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (55)..(9372)  
NAME/KEY: s1g\_peptide  
LOCATION: (55)..(120)  
US-09-562-702A-9

Query Match 84.0%; Score 16.8; DB 4; Length 9511;  
Best Local Similarity 90.0%; Pred. No. 53;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20  
DB 913 GGTCATGCCCCGGGCTTGTCC 9132

## RESULT 30

US-09-562-702A-9/c  
Sequence 9, Application US/09562702A  
Patent No. 6632790  
GENERAL INFORMATION:  
APPLICANT: Yurchenco, Peter  
TITLE OF INVENTION: Laminin 2 and Methods for Its Use  
FILE REFERENCE: 99-274-B  
CURRENT APPLICATION NUMBER: US/09/562,702A  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/155,945  
PRIOR FILING DATE: 1999-09-24  
PRIOR APPLICATION NUMBER: 60/143,289  
PRIOR FILING DATE: 1999-07-12  
PRIOR APPLICATION NUMBER: 60/139,198  
PRIOR FILING DATE: 1999-06-15  
PRIOR APPLICATION NUMBER: 60/131,720  
PRIOR FILING DATE: 1999-04-30  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatSeq Ver. 2.0  
SEQ ID NO 9  
LENGTH: 9511  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (55)..  
NAME/KEY: sig peptide  
LOCATION: (55)..  
US-09-562-702A-9

Query Match 84.0%; Score 16.8; DB 4; Length 9511;  
Best Local Similarity 90.0%; Pred. No. 53;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20  
DB 913 GGACATGCCCGGCGCATGACC 913

## RESULT 31

US-09-312-283C-355  
Sequence 355, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
FILE REFERENCE: 11000.1011c2  
CURRENT APPLICATION NUMBER: US/09/312,283C  
PRIOR FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 425  
SOFTWARE: PatSeq for Windows Version 4.0  
SEQ ID NO 355  
LENGTH: 473  
TYPE: DNA  
ORGANISM: Mouse  
US-09-312-283C-355

Query Match 77.0%; Score 15.4; DB 4; Length 473;  
Best Local Similarity 94.1%; Pred. No. 2.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATG 17  
DB 367 GGACATGCTCGGCATG 367

## RESULT 32

US-09-312-283C-355/c  
Sequence 355, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James G.  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
FILE REFERENCE: 11000.1011c2  
CURRENT APPLICATION NUMBER: US/09/312,283C  
PRIOR FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 425  
SOFTWARE: PatSeq for Windows Version 4.0  
SEQ ID NO 355  
LENGTH: 473  
TYPE: DNA  
ORGANISM: Mouse  
US-09-312-283C-355

Query Match 77.0%; Score 15.4; DB 4; Length 473;  
Best Local Similarity 94.1%; Pred. No. 2.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGCCCGGCGCATGTC 20  
DB 363 CATGCCCGGCGCATGTC 367

## RESULT 33

US-08-657-828A-3  
Sequence 3, Application US/08657828A  
Patent No. 5876711  
GENERAL INFORMATION:  
APPLICANT: Fattaey, Ali  
TITLE OF INVENTION: Methods and Compositions for Determining  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Onyx Pharmaceuticals, Inc.  
STREET: 3031 Research Drive  
CITY: Richmond  
STATE: CA  
COUNTRY: USA  
ZIP: 94806  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/657,828A  
FILING DATE: 31-MAY-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Glotta, Gregory  
REGISTRATION NUMBER: 32,028  
REFERENCE/DOCKET NUMBER: ONYX1021  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-262-8710

TELEFAX: 510-758-3405  
INFORMATION FOR SEQ ID NO: 3  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-657-828A-3

Query Match 76.0%; Score 15.2; DB 2; Length 20;  
Best Local Similarity 40.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20  
Db 1 RRCATGYRRRCATGYYY 20

## RESULT 34

US-08-657-828A-3/c  
Sequence 3, Application US/08657828A  
Patent No. 5876711  
GENERAL INFORMATION:  
APPLICANT: Fattaei, Ali  
TITLE OF INVENTION: Methods and Compositions for Determining  
TITLE OF INVENTION: the Tumor Suppressor Status of Cells  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Onyx Pharmaceuticals, Inc.  
STREET: 3031 Research Drive  
CITY: Richmond  
STATE: CA  
COUNTRY: USA  
ZIP: 94806  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/657,828A  
FILING DATE: 31-MAY-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gioceta, Gregory  
REGISTRATION NUMBER: 32,028  
REFERENCE/DOCKET NUMBER: ONYX1021  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-262-8710  
TELEFAX: 510-758-3405  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-657-828A-3

Query Match 76.0%; Score 15.2; DB 2; Length 20;  
Best Local Similarity 40.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20  
Db 20 RRCATGYRRRCATGYYY 1

## RESULT 35

US-09-260-420-3  
Sequence 3, Application US/09260420  
Patent No. 6391630  
GENERAL INFORMATION:  
APPLICANT: Fattaei, Ali  
TITLE OF INVENTION: Methods and Compositions for Determining  
TITLE OF INVENTION: the Tumor Suppressor Status of Cells  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Onyx Pharmaceuticals, Inc.  
STREET: 3031 Research Drive  
CITY: Richmond  
STATE: CA  
COUNTRY: USA  
ZIP: 94806  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/260,420  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/657,828  
FILING DATE: 31-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gioceta, Gregory  
REGISTRATION NUMBER: 32,028  
REFERENCE/DOCKET NUMBER: ONYX1021  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-262-8710  
TELEFAX: 510-758-3405  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-260-420-3

Query Match 76.0%; Score 15.2; DB 3; Length 20;  
Best Local Similarity 40.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20  
Db 1 RRCATGYRRRCATGYYY 20

## RESULT 36

US-09-260-420-3/c  
Sequence 3, Application US/09260420  
Patent No. 6391630  
GENERAL INFORMATION:  
APPLICANT: Fattaei, Ali  
TITLE OF INVENTION: Methods and Compositions for Determining  
TITLE OF INVENTION: the Tumor Suppressor Status of Cells  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Onyx Pharmaceuticals, Inc.  
STREET: 3031 Research Drive  
CITY: Richmond  
STATE: CA  
COUNTRY: USA  
ZIP: 94806  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

```

/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/260.420
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/657,828
/ FILING DATE: 31-MAY-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Giotta, Gregory
/ REGISTRATION NUMBER: 32,028
/ REFERENCE/DOCKET NUMBER: ONYX1021
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 510-262-8710
/ TELEFAX: 510-758-3405
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ US-09-260-420-3

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGCGCATGTC 20
DB      20 RRCATGYRRRCATGYT 1

```

```

RESULT 37
US-09-513-999C-22333
/ Sequence 22333, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ Patent No. 6783961
/ FILE REFERENCE: 59, US2, REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ CURRENT FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/122,487
/ PRIOR FILING DATE: 1999-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
/ SEQ ID NO 22333
/ LENGTH: 291
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 288
/ OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-22333

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 291;
Best Local Similarity 85.0%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGCGCATGTC 20
DB      160 GGACTTCCCGCGGATGTC 179

```

```

RESULT 38
US-09-513-999C-22333/c
/ Sequence 22333, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclet, A.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ Patent No. 6783961
/ FILE REFERENCE: 59, US2, REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ CURRENT FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/122,487
/ PRIOR FILING DATE: 1999-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
/ SEQ ID NO 22333
/ LENGTH: 291
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 288
/ OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-22333

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 291;
Best Local Similarity 85.0%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGCGCATGTC 20
DB      179 GGACATGCCCGGCGCATGTC 160

```

```

RESULT 39
US-09-828-000-1
/ Sequence 1, Application US/09828000
/ Patent No. 6596690
/ GENERAL INFORMATION:
/ APPLICANT: Government of the United States of America
/ TITLE OF INVENTION: Vasostatin as Marrow Protectant
/ FILE REFERENCE: 4239-55414
/ CURRENT APPLICATION NUMBER: US/09/828,000
/ CURRENT FILING DATE: 2001-04-06
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: Patent version 3.0
/ SEQ ID NO 1
/ LENGTH: 1251
/ TYPE: DNA
/ ORGANISM: Calreticulin
US-09-828-000-1

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 1251;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGCGCATGTC 20
DB      708 GGACATGCCCGGCGCATGTC 727

```

```

RESULT 40
US-09-828-000-1/c
/ Sequence 1, Application US/09828000
/ Patent No. 6596690
/ GENERAL INFORMATION:
/ APPLICANT: Government of the United States of America
/ TITLE OF INVENTION: Vasostatin as Marrow Protectant
/ FILE REFERENCE: 4239-55414
/ CURRENT APPLICATION NUMBER: US/09/828,000
/ CURRENT FILING DATE: 2001-04-06

```

NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 1  
LENGTH: 1251  
TYPE: DNA  
ORGANISM: Calreticulin  
US-09-828-000-1

Query Match 76.0%; Score 15.2; DB 4; Length 1251;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
Db 727 GGATATGCTCGGCGCTGTCC 708

RESULT 41  
US-09-906-393A-35  
Sequence 35, Application US/09906393A  
Patent No. 6780984  
GENERAL INFORMATION:  
APPLICANT: Wang, Zhou  
APPLICANT: Xiao, Wuhan  
TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED  
FILE REFERENCE: 1720-1-001CIP  
CURRENT APPLICATION NUMBER: US/09/906,393A  
CURRENT FILING DATE: 2001-07-16  
PRIOR FILING DATE: 2000-07-17  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 35  
LENGTH: 1920  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-906-393A-35

Query Match 76.0%; Score 15.2; DB 4; Length 1920;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
Db 778 GGACAGCCCGAGCATATCC 797

RESULT 42  
US-09-906-393A-35/c  
Sequence 35, Application US/09906393A  
Patent No. 6780984  
GENERAL INFORMATION:  
APPLICANT: Wang, Zhou  
APPLICANT: Xiao, Wuhan  
TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED  
FILE REFERENCE: 1720-1-001CIP  
CURRENT APPLICATION NUMBER: US/09/906,393A  
CURRENT FILING DATE: 2001-07-16  
PRIOR FILING DATE: 2000-07-17  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 35  
LENGTH: 1920  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-906-393A-35

Query Match 76.0%; Score 15.2; DB 4; Length 1920;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20

Db 797 GGATATGCTCGGCGCTGTCC 778

RESULT 43  
US-09-702-327-3  
Sequence 3, Application US/09702327  
Patent No. 6426220  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Lex M. Cowsett  
TITLE OF INVENTION: ANTISENSE MODULATION OF CALRETICULIN EXPRESSION  
FILE REFERENCE: RTS-0097  
CURRENT APPLICATION NUMBER: US/09/702,327  
CURRENT FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 3  
LENGTH: 1958  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (109)..(1362)  
US-09-702-327-3

Query Match 76.0%; Score 15.2; DB 4; Length 1958;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
Db 816 GGACAGCCCGAGCATATCC 835

RESULT 44  
US-09-702-327-3/c  
Sequence 3, Application US/09702327  
Patent No. 6426220  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Lex M. Cowsett  
TITLE OF INVENTION: ANTISENSE MODULATION OF CALRETICULIN EXPRESSION  
FILE REFERENCE: RTS-0097  
CURRENT APPLICATION NUMBER: US/09/702,327  
CURRENT FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 3  
LENGTH: 1958  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (109)..(1362)  
US-09-702-327-3

Query Match 76.0%; Score 15.2; DB 4; Length 1958;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
Db 835 GGATATGCTCGGCGCTGTCC 816

RESULT 45  
US-08-460-309-3  
Sequence 3, Application US/08460309  
Patent No. 5837496  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leivo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin Fragments and Uses Thereof

NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,309  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/125,077  
FILING DATE: 22-SEP-1993  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-309-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20  
DB 920 GGTCTATGCCAGGCGCTTGTC 939

RESULT 46  
US-08-460-309-3/c  
Sequence 3, Application US/08460309  
Patent No. 5837496  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leiyo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,309  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/125,077  
FILING DATE: 22-SEP-1993  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-309-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20  
DB 939 GGACATGCCCGGCGCATGCC 920

RESULT 47  
US-08-125-077-3  
Sequence 3, Application US/08125077  
Patent No. 5872231  
Patent No. 5872231 5840863  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leiyo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/125,077  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-125-077-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 920 GGTCATGCCAGGCGCTGTGCC 939

RESULT 48  
US-08-125-077-3/c  
Sequence 3, Application US/08125077  
Patent No. 5872231  
Patent No. 5872231 5840863  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leiyo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/125,077  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-125-077-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 939 GGACAGCCCTGGCATGACC 920

RESULT 49  
US-09-562-702A-7  
Sequence 7, Application US/09562702A  
Patent No. 6632790  
GENERAL INFORMATION:  
APPLICANT: Yurchenco, Peter  
TITLE OF INVENTION: Laminin 2 and Methods for Its Use  
FILE REFERENCE: 99-274-B  
CURRENT FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/155,945  
PRIOR FILING DATE: 1999-09-24  
PRIOR APPLICATION NUMBER: 60/143,289  
PRIOR FILING DATE: 1999-07-12  
PRIOR APPLICATION NUMBER: 60/139,198  
PRIOR FILING DATE: 1999-06-15  
PRIOR APPLICATION NUMBER: 60/131,720  
PRIOR FILING DATE: 1999-04-30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7

Query Match 76.0%; Score 15.2; DB 4; Length 9419;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20  
DB 805 GGTCATGCCAGGCGCTGTGCC 824

RESULT 50  
US-09-562-702A-7/c  
Sequence 7, Application US/09562702A  
Patent No. 6632790  
GENERAL INFORMATION:  
APPLICANT: Yurchenco, Peter  
TITLE OF INVENTION: Laminin 2 and Methods for Its Use  
FILE REFERENCE: 99-274-B  
CURRENT FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/155,945  
PRIOR FILING DATE: 1999-09-24  
PRIOR APPLICATION NUMBER: 60/143,289  
PRIOR FILING DATE: 1999-07-12  
PRIOR APPLICATION NUMBER: 60/139,198  
PRIOR FILING DATE: 1999-06-15  
PRIOR APPLICATION NUMBER: 60/131,720  
PRIOR FILING DATE: 1999-04-30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7

```

; LENGTH: 9419
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(9264)
US-09-562-702A-7

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 9419;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGGCATGTCC 20
          ||||| ||||| ||||| ||
Db      824 GGACAGCCCTGGCATGACC 805

```

Search completed: January 14, 2005, 17:39:57  
 Job time : 88.0526 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 16:14:28 ; Search time 348.421 Seconds  
(without alignments)  
329.824 Million cell updates/sec

Title: US-09-578-453-2

Perfect score: 20

Sequence: 1 GCAGATCCCGGCGATGTC 20

Scoring table: IDENTITY NUC  
Gapc 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 500 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCRT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCRTS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	16	US-10-447-136-162 Sequence 162, App
2	20	100.0	20	16	US-10-447-136-162 Sequence 162, App
3	20	100.0	24	14	US-10-023-318-19 Sequence 19, App1
4	20	100.0	24	14	US-10-023-318-19 Sequence 19, App1
5	20	100.0	24	14	US-10-023-318-38 Sequence 38, App1
6	20	100.0	24	14	US-10-023-318-38 Sequence 38, App1
7	19	95.0	19	15	US-10-339-161-14 Sequence 14, App1
8	19	95.0	19	15	US-10-339-161-14 Sequence 14, App1
9	17.4	87.0	729	17	US-10-767-701-2482 Sequence 2482, Ap
10	17.4	87.0	729	17	US-10-767-701-2482 Sequence 2482, Ap
11	17.4	87.0	3325	16	US-10-425-114-30682 Sequence 30682, A
12	17.4	87.0	3325	16	US-10-425-114-30682 Sequence 30682, A

13	17.4	87.0	3458	18	US-10-425-115-4271 Sequence 4271, Ap
14	17.4	87.0	3458	18	US-10-425-115-4271 Sequence 4271, Ap
15	17.4	87.0	3569	16	US-10-425-114-2765 Sequence 2765, Ap
16	17.4	87.0	3569	16	US-10-425-114-2765 Sequence 2765, Ap
17	17.4	87.0	3609	18	US-10-425-115-4270 Sequence 4270, Ap
18	17.4	87.0	3609	18	US-10-425-115-4270 Sequence 4270, Ap
19	16.8	84.0	30	10	US-09-829-922-31 Sequence 31, App1
20	16.8	84.0	30	10	US-09-829-922-31 Sequence 31, App1
21	16.8	84.0	30	15	US-10-160-290-31 Sequence 31, App1
22	16.8	84.0	30	15	US-10-160-290-31 Sequence 31, App1
23	16.8	84.0	21565	16	US-10-052-482-61 Sequence 61, App1
24	16.8	84.0	21565	16	US-10-052-482-61 Sequence 61, App1
25	16.8	84.0	46030	17	US-10-367-094-1 Sequence 1, App1
26	16.8	84.0	46030	17	US-10-367-094-1 Sequence 1, App1
27	16.8	84.0	60327	16	US-10-052-482-187 Sequence 187, App
28	16.8	84.0	60327	16	US-10-052-482-187 Sequence 187, App
29	16.8	84.0	68233	15	US-10-034-650-31 Sequence 31, App1
30	16.8	84.0	68233	15	US-10-034-650-31 Sequence 31, App1
31	16.8	84.0	77941	13	US-10-087-192-709 Sequence 709, App
32	16.8	84.0	77941	13	US-10-087-192-709 Sequence 709, App
33	16.8	84.0	145068	17	US-10-322-281-33 Sequence 33, App1
34	16.8	84.0	145068	17	US-10-322-281-33 Sequence 33, App1
35	16.4	82.0	400	10	US-09-835-9768-77 Sequence 77, App1
36	16.4	82.0	400	10	US-09-835-9768-77 Sequence 77, App1
37	16.4	82.0	3422	10	US-09-835-9768-111 Sequence 111, App
38	16.4	82.0	3422	10	US-09-835-9768-111 Sequence 111, App
39	16.4	82.0	32351	17	US-10-322-696-40 Sequence 40, App1
40	16.4	82.0	32351	17	US-10-322-696-40 Sequence 40, App1
41	16.4	82.0	203264	13	US-10-087-192-988 Sequence 988, App
42	16.4	82.0	203264	13	US-10-087-192-988 Sequence 988, App
43	15.8	79.0	274	9	US-09-294-0938-1681 Sequence 1681, Ap
44	15.8	79.0	274	9	US-09-294-0938-1681 Sequence 1681, Ap
45	15.8	79.0	277	10	US-09-294-0938-580 Sequence 580, App
46	15.8	79.0	277	10	US-09-294-0938-580 Sequence 580, App
47	15.8	79.0	294	9	US-09-294-0938-5575 Sequence 5575, Ap
48	15.8	79.0	294	9	US-09-294-0938-5575 Sequence 5575, Ap
49	15.8	79.0	482	9	US-09-783-590-4831 Sequence 4831, Ap
50	15.8	79.0	482	9	US-09-783-590-4831 Sequence 4831, Ap
51	15.8	79.0	501	16	US-10-282-1228-17485 Sequence 17485, A
52	15.8	79.0	501	16	US-10-282-1228-17485 Sequence 17485, A
53	15.8	79.0	1217	16	US-10-425-114-20194 Sequence 20194, A
54	15.8	79.0	1217	16	US-10-425-114-20194 Sequence 20194, A
55	15.8	79.0	1539	15	US-10-156-761-998 Sequence 998, App
56	15.8	79.0	1539	15	US-10-156-761-998 Sequence 998, App
57	15.8	79.0	2253	18	US-10-425-115-141130 Sequence 141130, Sequence 141130,
58	15.8	79.0	2253	18	US-10-425-115-141130 Sequence 141130, Sequence 141130,
59	15.8	79.0	330973	13	US-10-087-192-1498 Sequence 1498, Ap
60	15.8	79.0	330973	13	US-10-087-192-1498 Sequence 1498, Ap
61	15.8	79.0	438892	13	US-10-087-192-454 Sequence 454, App
62	15.8	79.0	438892	13	US-10-087-192-454 Sequence 454, App
63	15.8	79.0	717651	18	US-10-719-993-6817 Sequence 6817, Ap
64	15.8	79.0	717651	18	US-10-719-993-6817 Sequence 6817, Ap
65	15.8	79.0	2731748	17	US-10-297-465A-1 Sequence 1, App1
66	15.8	79.0	2731748	17	US-10-297-465A-1 Sequence 1, App1
67	15.8	79.0	9025608	15	US-10-156-761-1 Sequence 1, App1
68	15.8	79.0	9025608	15	US-10-156-761-1 Sequence 1, App1
69	15.4	77.0	473	10	US-09-866-050A-355 Sequence 355, App
70	15.4	77.0	473	10	US-09-866-050A-355 Sequence 355, App
71	15.4	77.0	473	14	US-10-152-661-355 Sequence 355, App
72	15.4	77.0	473	14	US-10-152-661-355 Sequence 355, App
73	15.4	77.0	764	13	US-10-027-632-142160 Sequence 142160, Sequence 142160,
74	15.4	77.0	764	13	US-10-027-632-142160 Sequence 142160, Sequence 142160,
75	15.4	77.0	764	15	US-10-027-632-142160 Sequence 142160, Sequence 142160,
76	15.4	77.0	764	15	US-10-027-632-142160 Sequence 142160, Sequence 142160,
77	15.4	77.0	1701	15	US-10-369-493-39872 Sequence 39872, A
78	15.4	77.0	1701	15	US-10-369-493-39872 Sequence 39872, A
79	15.4	77.0	1857	15	US-10-369-493-39490 Sequence 39490, A
80	15.4	77.0	1857	15	US-10-369-493-39490 Sequence 39490, A
81	15.4	77.0	2315	15	US-10-369-493-39122 Sequence 39122, A
82	15.4	77.0	2315	15	US-10-369-493-39122 Sequence 39122, A
83	15.4	77.0	7282	15	US-10-213-948-11 Sequence 11, App1
84	15.4	77.0	7282	15	US-10-213-948-11 Sequence 11, App1
85	15.4	77.0	57013	13	US-10-087-192-1798 Sequence 1798, Ap

C 86	15.4	77.0	57013	13	US-10-087-192-1798	Sequence 1798, Ap
C 87	15.4	77.0	114411	18	US-10-723-860-3910	Sequence 3910, Ap
C 88	15.4	77.0	114411	18	US-10-723-860-3910	Sequence 3910, Ap
C 89	15.2	76.0	297	9	US-09-878-722-72	Sequence 72, Appl
C 90	15.2	76.0	297	9	US-09-878-722-72	Sequence 72, Appl
C 91	15.2	76.0	297	9	US-09-878-178-1454	Sequence 1454, Ap
C 92	15.2	76.0	297	9	US-09-878-178-1454	Sequence 1454, Ap
C 93	15.2	76.0	297	9	US-09-904-456-72	Sequence 72, Appl
C 94	15.2	76.0	297	10	US-09-904-456-72	Sequence 72, Appl
C 95	15.2	76.0	297	13	US-10-046-935-1454	Sequence 1454, Ap
C 96	15.2	76.0	297	13	US-10-046-935-1454	Sequence 1454, Ap
C 97	15.2	76.0	297	14	US-10-146-502-1454	Sequence 1454, Ap
C 98	15.2	76.0	297	14	US-10-146-502-1454	Sequence 1454, Ap
C 99	15.2	76.0	298	18	US-10-425-115-39745	Sequence 39745, A
C 100	15.2	76.0	298	18	US-10-425-115-39745	Sequence 39745, A
C 101	15.2	76.0	366	18	US-10-744-936-4	Sequence 4, Appl1
C 102	15.2	76.0	366	18	US-10-744-936-4	Sequence 4, Appl1
C 103	15.2	76.0	401	10	US-09-918-995-4352	Sequence 4352, Ap
C 104	15.2	76.0	401	10	US-09-918-995-4352	Sequence 4352, Ap
C 105	15.2	76.0	456	16	US-10-424-599-60443	Sequence 60443, A
C 106	15.2	76.0	456	16	US-10-424-599-60443	Sequence 60443, A
C 107	15.2	76.0	472	10	US-09-918-995-13145	Sequence 13145, A
C 108	15.2	76.0	472	10	US-09-918-995-13145	Sequence 13145, A
C 109	15.2	76.0	491	10	US-09-918-995-11130	Sequence 11130, A
C 110	15.2	76.0	491	10	US-09-918-995-11130	Sequence 11130, A
C 111	15.2	76.0	492	10	US-09-918-995-1921	Sequence 1921, Ap
C 112	15.2	76.0	492	10	US-10-156-761-1173	Sequence 1173, Ap
C 113	15.2	76.0	498	15	US-10-156-761-1173	Sequence 1173, Ap
C 114	15.2	76.0	498	15	US-10-156-761-1173	Sequence 1173, Ap
C 115	15.2	76.0	516	16	US-10-424-599-106147	Sequence 106147, A
C 116	15.2	76.0	516	16	US-10-424-599-106147	Sequence 106147, A
C 117	15.2	76.0	538	15	US-10-029-386-13284	Sequence 13284, A
C 118	15.2	76.0	538	15	US-10-029-386-13284	Sequence 13284, A
C 119	15.2	76.0	569	15	US-10-029-386-13639	Sequence 13639, A
C 120	15.2	76.0	569	15	US-10-029-386-13639	Sequence 13639, A
C 121	15.2	76.0	574	15	US-10-029-386-7058	Sequence 7058, Ap
C 122	15.2	76.0	574	15	US-10-029-386-7058	Sequence 7058, Ap
C 123	15.2	76.0	586	18	US-10-425-115-78128	Sequence 78128, A
C 124	15.2	76.0	586	18	US-10-425-115-78128	Sequence 78128, A
C 125	15.2	76.0	615	15	US-10-163-198-32	Sequence 32, Appl
C 126	15.2	76.0	615	15	US-10-163-198-32	Sequence 32, Appl
C 127	15.2	76.0	627	16	US-10-424-599-104831	Sequence 104831, A
C 128	15.2	76.0	627	16	US-10-424-599-104831	Sequence 104831, A
C 129	15.2	76.0	1008	15	US-10-156-761-2985	Sequence 2985, Ap
C 130	15.2	76.0	1008	15	US-10-156-761-2985	Sequence 2985, Ap
C 131	15.2	76.0	1152	15	US-10-259-165-201	Sequence 201, App
C 132	15.2	76.0	1152	15	US-10-259-165-201	Sequence 201, App
C 133	15.2	76.0	1251	10	US-09-828-000-1	Sequence 1, Appl1
C 134	15.2	76.0	1251	10	US-09-828-000-1	Sequence 1, Appl1
C 135	15.2	76.0	1251	15	US-10-405-588-1	Sequence 1, Appl1
C 136	15.2	76.0	1251	15	US-10-405-588-1	Sequence 1, Appl1
C 137	15.2	76.0	1308	15	US-10-156-761-752	Sequence 752, App
C 138	15.2	76.0	1308	15	US-10-156-761-752	Sequence 752, App
C 139	15.2	76.0	1513	16	US-10-424-599-104834	Sequence 104834, A
C 140	15.2	76.0	1513	16	US-10-424-599-104834	Sequence 104834, A
C 141	15.2	76.0	1569	17	US-10-437-963-53491	Sequence 53491, A
C 142	15.2	76.0	1569	17	US-10-437-963-53491	Sequence 53491, A
C 143	15.2	76.0	1716	13	US-10-027-632-24965	Sequence 24965, A
C 144	15.2	76.0	1716	13	US-10-027-632-24965	Sequence 24965, A
C 145	15.2	76.0	1716	15	US-10-027-632-24965	Sequence 24965, A
C 146	15.2	76.0	1899	16	US-10-439-703-7	Sequence 7, Appl1
C 147	15.2	76.0	1899	16	US-10-439-703-7	Sequence 7, Appl1
C 148	15.2	76.0	1920	10	US-09-906-393A-35	Sequence 35, Appl
C 149	15.2	76.0	1920	10	US-09-906-393A-35	Sequence 35, Appl
C 150	15.2	76.0	1924	15	US-10-240-965-265	Sequence 265, App
C 151	15.2	76.0	1924	15	US-10-240-965-265	Sequence 265, App
C 152	15.2	76.0	1940	15	US-10-367-093-13	Sequence 13, Appl
C 153	15.2	76.0	1940	15	US-10-367-093-13	Sequence 13, Appl
C 154	15.2	76.0	1958	9	US-09-791-406-3	Sequence 3, Appl1
C 155	15.2	76.0	1958	9	US-09-791-406-3	Sequence 3, Appl1
C 156	15.2	76.0	3078	13	US-10-027-632-177166	Sequence 177166, A
C 157	15.2	76.0	3078	13	US-10-027-632-177166	Sequence 177166, A
C 158	15.2	76.0	3078	13	US-10-027-632-177166	Sequence 177166, A
C 159	15.2	76.0	3078	15	US-10-027-632-177166	Sequence 177166, A
C 160	15.2	76.0	3078	15	US-10-027-632-177166	Sequence 177166, A
C 161	15.2	76.0	4258	15	US-10-128-714-494	Sequence 494, App
C 162	15.2	76.0	4258	15	US-10-128-714-494	Sequence 494, App
C 163	15.2	76.0	4258	15	US-10-128-714-544	Sequence 544, App
C 164	15.2	76.0	4258	15	US-10-128-714-544	Sequence 544, App
C 165	15.2	76.0	5340	17	US-10-437-963-9004	Sequence 9004, Ap
C 166	15.2	76.0	5340	17	US-10-437-963-9004	Sequence 9004, Ap
C 167	15.2	76.0	9534	9	US-09-954-531-1380	Sequence 1380, Ap
C 168	15.2	76.0	9534	9	US-09-954-531-1380	Sequence 1380, Ap
C 169	15.2	76.0	9649	18	US-10-723-860-697	Sequence 697, App
C 170	15.2	76.0	9649	18	US-10-723-860-697	Sequence 697, App
C 171	15.2	76.0	9670	18	US-10-723-860-5324	Sequence 5324, Ap
C 172	15.2	76.0	9670	18	US-10-723-860-5324	Sequence 5324, Ap
C 173	15.2	76.0	10867	9	US-09-764-869-2113	Sequence 2113, Ap
C 174	15.2	76.0	10867	9	US-09-764-869-2113	Sequence 2113, Ap
C 175	15.2	76.0	10867	9	US-09-764-869-3970	Sequence 3970, Ap
C 176	15.2	76.0	10867	9	US-09-764-869-3970	Sequence 3970, Ap
C 177	15.2	76.0	10867	14	US-10-091-504-2113	Sequence 2113, Ap
C 178	15.2	76.0	10867	14	US-10-091-504-2113	Sequence 2113, Ap
C 179	15.2	76.0	10867	16	US-10-227-577-2113	Sequence 2113, Ap
C 180	15.2	76.0	10867	16	US-10-227-577-2113	Sequence 2113, Ap
C 181	15.2	76.0	10867	16	US-10-242-515-3970	Sequence 3970, Ap
C 182	15.2	76.0	10867	16	US-10-242-515-3970	Sequence 3970, Ap
C 183	15.2	76.0	10894	9	US-09-764-869-2111	Sequence 2111, Ap
C 184	15.2	76.0	10894	9	US-09-764-869-2111	Sequence 2111, Ap
C 185	15.2	76.0	10894	14	US-10-091-504-2111	Sequence 2111, Ap
C 186	15.2	76.0	10894	14	US-10-091-504-2111	Sequence 2111, Ap
C 187	15.2	76.0	10894	16	US-10-227-577-2111	Sequence 2111, Ap
C 188	15.2	76.0	10894	16	US-10-227-577-2111	Sequence 2111, Ap
C 189	15.2	76.0	10907	9	US-09-764-869-2112	Sequence 2112, Ap
C 190	15.2	76.0	10907	9	US-09-764-869-2112	Sequence 2112, Ap
C 191	15.2	76.0	10907	9	US-09-764-877-3968	Sequence 3968, Ap
C 192	15.2	76.0	10907	9	US-09-764-877-3968	Sequence 3968, Ap
C 193	15.2	76.0	10907	14	US-10-091-504-2112	Sequence 2112, Ap
C 194	15.2	76.0	10907	14	US-10-091-504-2112	Sequence 2112, Ap
C 195	15.2	76.0	10907	16	US-10-227-577-2112	Sequence 2112, Ap
C 196	15.2	76.0	10907	16	US-10-227-577-2112	Sequence 2112, Ap
C 197	15.2	76.0	10907	16	US-10-242-515-3968	Sequence 3968, Ap
C 198	15.2	76.0	10907	16	US-10-242-515-3968	Sequence 3968, Ap
C 199	15.2	76.0	12718	9	US-09-764-877-3972	Sequence 3972, Ap
C 200	15.2	76.0	12718	9	US-09-764-877-3972	Sequence 3972, Ap
C 201	15.2	76.0	12718	16	US-10-242-515-3972	Sequence 3972, Ap
C 202	15.2	76.0	12718	16	US-10-242-515-3972	Sequence 3972, Ap
C 203	15.2	76.0	33097	13	US-10-087-192-241	Sequence 241, App
C 204	15.2	76.0	33097	13	US-10-087-192-241	Sequence 241, App
C 205	15.2	76.0	39405	13	US-10-087-192-1285	Sequence 1285, App
C 206	15.2	76.0	39405	13	US-10-087-192-1285	Sequence 1285, App
C 207	15.2	76.0	50720	17	US-10-322-281-621	Sequence 621, App
C 208	15.2	76.0	50720	17	US-10-322-281-621	Sequence 621, App
C 209	15.2	76.0	67191	11	US-09-997-722-169	Sequence 169, App
C 210	15.2	76.0	67191	11	US-09-997-722-169	Sequence 169, App
C 211	15.2	76.0	67191	15	US-10-105-612-1	Sequence 1, Appl1
C 212	15.2	76.0	67191	15	US-10-105-612-1	Sequence 1, Appl1
C 213	15.2	76.0	67191	15	US-10-105-612-1	Sequence 1, Appl1
C 214	15.2	76.0	110096	9	US-09-880-107-1542	Sequence 1542, Ap
C 215	15.2	76.0	110096	9	US-09-880-107-1542	Sequence 1542, Ap
C 216	15.2	76.0	163701	17	US-10-322-281-439	Sequence 439, App
C 217	15.2	76.0	163701	17	US-10-322-281-439	Sequence 439, App
C 218	15.2	76.0	173808	13	US-10-003-806-10	Sequence 10, Appl
C 219	15.2	76.0	173808	13	US-10-003-806-10	Sequence 10, Appl
C 220	15.2	76.0	347001	17	US-10-319-908-16	Sequence 16, Appl
C 221	15.2	76.0	347001	17	US-10-319-908-16	Sequence 16, Appl
C 222	14.8	74.0	49	9	US-09-813-824A-31	Sequence 31, Appl
C 223	14.8	74.0	201	17	US-10-741-601-14449	Sequence 14449, A
C 224	14.8	74.0	201	17	US-10-741-601-14449	Sequence 14449, A
C 225	14.8	74.0	400	16	US-10-424-599-90906	Sequence 90906, A
C 226	14.8	74.0	400	16	US-10-424-599-90906	Sequence 90906, A
C 227	14.8	74.0	478	10	US-09-918-995-23256	Sequence 23256, A
C 228	14.8	74.0	478	10	US-09-918-995-23256	Sequence 23256, A
C 229	14.8	74.0	566	9	US-09-920-300A-647	Sequence 647, App
C 230	14.8	74.0	566	9	US-09-920-300A-647	Sequence 647, App
C 231	14.8	74.0	566	13	US-10-033-528-647	Sequence 647, App

C 232	14.8	74.0	566	13	US-10-033-528-647	Sequence 647, App	305	14.4	72.0	792	13	US-10-027-632-14097	Sequence 14097, A
C 233	14.8	74.0	566	15	US-10-099-926-647	Sequence 647, App	306	14.4	72.0	792	13	US-10-027-632-14097	Sequence 14097, A
C 234	14.8	74.0	566	15	US-10-099-926-647	Sequence 647, App	307	14.4	72.0	792	13	US-10-027-632-14098	Sequence 14098, A
C 235	14.8	74.0	747	9	US-09-813-453A-35	Sequence 35, App1	308	14.4	72.0	792	13	US-10-027-632-14098	Sequence 14098, A
C 236	14.8	74.0	747	9	US-09-813-453A-35	Sequence 35, App1	309	14.4	72.0	792	13	US-10-027-632-14099	Sequence 14099, A
C 237	14.8	74.0	789	13	US-10-027-632-12458	Sequence 12458, A	310	14.4	72.0	792	13	US-10-027-632-14099	Sequence 14099, A
C 238	14.8	74.0	789	13	US-10-027-632-12458	Sequence 12458, A	311	14.4	72.0	792	15	US-10-027-632-14096	Sequence 14096, A
C 239	14.8	74.0	789	15	US-10-027-632-12458	Sequence 12458, A	312	14.4	72.0	792	15	US-10-027-632-14096	Sequence 14096, A
C 240	14.8	74.0	789	15	US-10-027-632-12458	Sequence 12458, A	313	14.4	72.0	792	15	US-10-027-632-14097	Sequence 14097, A
C 241	14.8	74.0	811	13	US-10-159-749-5	Sequence 5, App11	314	14.4	72.0	792	15	US-10-027-632-14097	Sequence 14097, A
C 242	14.8	74.0	811	13	US-10-159-749-5	Sequence 5, App11	315	14.4	72.0	792	15	US-10-027-632-14098	Sequence 14098, A
C 243	14.8	74.0	994	18	US-10-653-047-30	Sequence 30, App1	316	14.4	72.0	792	15	US-10-027-632-14098	Sequence 14098, A
C 244	14.8	74.0	994	18	US-10-653-047-30	Sequence 30, App1	317	14.4	72.0	792	15	US-10-027-632-14099	Sequence 14099, A
C 245	14.8	74.0	1029	18	US-10-425-115-95380	Sequence 95380, A	318	14.4	72.0	792	15	US-10-027-632-14099	Sequence 14099, A
C 246	14.8	74.0	1029	18	US-10-425-115-95380	Sequence 95380, A	319	14.4	72.0	1150	16	US-10-260-238-3780	Sequence 3780, App
C 247	14.8	74.0	1057	18	US-10-425-115-14357	Sequence 14357, A	320	14.4	72.0	1150	16	US-10-260-238-3780	Sequence 3780, App
C 248	14.8	74.0	1057	18	US-10-425-115-14357	Sequence 14357, A	321	14.4	72.0	1611	15	US-10-369-493-44284	Sequence 44284, A
C 249	14.8	74.0	1354	17	US-10-437-963-71006	Sequence 71006, A	322	14.4	72.0	1611	15	US-10-369-493-44284	Sequence 44284, A
C 250	14.8	74.0	1354	17	US-10-437-963-71006	Sequence 71006, A	323	14.4	72.0	1651	17	US-10-767-701-13956	Sequence 13956, A
C 251	14.8	74.0	1725	17	US-10-437-963-60963	Sequence 60963, A	324	14.4	72.0	1651	17	US-10-767-701-13956	Sequence 13956, A
C 252	14.8	74.0	1725	17	US-10-437-963-60963	Sequence 60963, A	325	14.4	72.0	1677	15	US-10-369-493-33181	Sequence 33181, A
C 253	14.8	74.0	1781	15	US-10-316-754-20	Sequence 20, App1	326	14.4	72.0	1677	15	US-10-369-493-33181	Sequence 33181, A
C 254	14.8	74.0	1781	15	US-10-316-754-20	Sequence 20, App1	327	14.4	72.0	2002	14	US-10-153-668-173	Sequence 173, App
C 255	14.8	74.0	1949	15	US-10-316-754-12	Sequence 12, App1	328	14.4	72.0	2002	14	US-10-153-668-173	Sequence 173, App
C 256	14.8	74.0	1949	15	US-10-316-754-12	Sequence 12, App1	329	14.4	72.0	3744	14	US-10-153-668-175	Sequence 175, App
C 257	14.8	74.0	2044	16	US-10-425-114-31320	Sequence 31320, A	330	14.4	72.0	3744	14	US-10-153-668-175	Sequence 175, App
C 258	14.8	74.0	2044	16	US-10-425-114-31320	Sequence 31320, A	331	14.4	72.0	3901	14	US-10-153-668-175	Sequence 175, App
C 259	14.8	74.0	2087	16	US-10-425-114-4515	Sequence 4515, App	332	14.4	72.0	3901	14	US-10-153-668-175	Sequence 175, App
C 260	14.8	74.0	2087	16	US-10-425-114-4515	Sequence 4515, App	333	14.4	72.0	5066	10	US-09-824-574-1	Sequence 1, App11
C 261	14.8	74.0	2088	15	US-10-156-761-1788	Sequence 1788, App	334	14.4	72.0	5066	10	US-09-824-574-1	Sequence 1, App11
C 262	14.8	74.0	2152	18	US-10-425-115-95382	Sequence 95382, A	335	14.4	72.0	30000	12	US-09-980-217-2	Sequence 2, App11
C 263	14.8	74.0	2152	18	US-10-425-115-95382	Sequence 95382, A	336	14.4	72.0	30000	12	US-09-980-217-2	Sequence 2, App11
C 264	14.8	74.0	2152	18	US-10-425-115-95382	Sequence 95382, A	337	14.4	72.0	40394	17	US-10-741-601-5574	Sequence 5774, App
C 265	14.8	74.0	3621	17	US-10-437-963-12579	Sequence 12579, A	338	14.4	72.0	40394	17	US-10-741-601-5574	Sequence 5774, App
C 266	14.8	74.0	3621	17	US-10-437-963-12579	Sequence 12579, A	339	14.4	72.0	59914	17	US-10-741-601-55619	Sequence 5619, App
C 267	14.8	74.0	3934	9	US-09-813-453A-77	Sequence 77, App1	340	14.4	72.0	59914	17	US-10-741-601-55619	Sequence 5619, App
C 268	14.8	74.0	3934	9	US-09-813-453A-77	Sequence 77, App1	341	14.4	72.0	101782	17	US-10-741-601-55619	Sequence 5661, App
C 269	14.8	74.0	63693	17	US-10-741-601-5560	Sequence 5560, App	342	14.4	72.0	101782	17	US-10-741-601-55619	Sequence 5661, App
C 270	14.8	74.0	63693	17	US-10-741-601-5560	Sequence 5560, App	343	14.4	72.0	157044	13	US-10-087-192-565	Sequence 565, App
C 271	14.8	74.0	87350	16	US-10-374-077-79	Sequence 79, App1	344	14.4	72.0	157044	13	US-10-087-192-565	Sequence 565, App
C 272	14.8	74.0	87350	16	US-10-374-077-79	Sequence 79, App1	345	14.4	72.0	2256646	17	US-10-470-565-1	Sequence 1, App11
C 273	14.8	74.0	151870	17	US-10-741-601-5614	Sequence 5614, App	346	14.4	72.0	2256646	17	US-10-470-565-1	Sequence 1, App11
C 274	14.8	74.0	151870	17	US-10-741-601-5614	Sequence 5614, App	347	14.2	71.0	183	17	US-10-437-963-100008	Sequence 100008, A
C 275	14.8	74.0	313287	17	US-10-322-281-48	Sequence 48, App1	348	14.2	71.0	183	17	US-10-437-963-100008	Sequence 100008, A
C 276	14.8	74.0	313287	17	US-10-322-281-48	Sequence 48, App1	349	14.2	71.0	201	18	US-10-719-993-41925	Sequence 41925, A
C 277	14.8	74.0	1601042	13	US-10-027-632-55064	Sequence 55064, A	350	14.2	71.0	201	18	US-10-719-993-41925	Sequence 41925, A
C 278	14.8	74.0	1601042	13	US-10-027-632-55064	Sequence 55064, A	351	14.2	71.0	228	18	US-10-425-115-116704	Sequence 116704, A
C 279	14.8	74.0	1601042	15	US-10-027-632-55064	Sequence 55064, A	352	14.2	71.0	228	18	US-10-425-115-116704	Sequence 116704, A
C 280	14.8	74.0	1601042	15	US-10-027-632-55064	Sequence 55064, A	353	14.2	71.0	237	17	US-10-469-285-383	Sequence 383, App
C 281	14.4	72.0	187	18	US-10-653-047-7853	Sequence 7853, App	354	14.2	71.0	237	17	US-10-469-285-383	Sequence 383, App
C 282	14.4	72.0	187	18	US-10-653-047-7853	Sequence 7853, App	355	14.2	71.0	250	11	US-09-864-408A-4661	Sequence 4661, App
C 283	14.4	72.0	200	18	US-10-425-115-159810	Sequence 159810, A	356	14.2	71.0	250	11	US-09-864-408A-4661	Sequence 4661, App
C 284	14.4	72.0	200	18	US-10-425-115-159810	Sequence 159810, A	357	14.2	71.0	261	15	US-10-245-618-9	Sequence 9, App11
C 285	14.4	72.0	352	18	US-10-425-115-87386	Sequence 87386, A	358	14.2	71.0	261	15	US-10-245-618-9	Sequence 9, App11
C 286	14.4	72.0	352	18	US-10-425-115-87386	Sequence 87386, A	359	14.2	71.0	261	15	US-10-245-618-9	Sequence 9, App11
C 287	14.4	72.0	370	15	US-10-106-698-2499	Sequence 2499, App	360	14.2	71.0	261	15	US-10-245-618-9	Sequence 9, App11
C 288	14.4	72.0	370	15	US-10-106-698-2499	Sequence 2499, App	361	14.2	71.0	294	18	US-10-425-115-154546	Sequence 154546, A
C 289	14.4	72.0	399	18	US-10-425-115-130668	Sequence 130668, A	362	14.2	71.0	294	18	US-10-425-115-154546	Sequence 154546, A
C 290	14.4	72.0	399	18	US-10-425-115-130668	Sequence 130668, A	363	14.2	71.0	360	16	US-10-282-122A-19545	Sequence 19525, A
C 291	14.4	72.0	453	18	US-10-425-115-128493	Sequence 128493, A	364	14.2	71.0	360	16	US-10-282-122A-19545	Sequence 19525, A
C 292	14.4	72.0	453	18	US-10-425-115-128493	Sequence 128493, A	365	14.2	71.0	375	17	US-10-437-963-10573	Sequence 10573, A
C 293	14.4	72.0	473	16	US-10-260-238-3779	Sequence 3779, App	366	14.2	71.0	375	17	US-10-437-963-10573	Sequence 10573, A
C 294	14.4	72.0	473	16	US-10-260-238-3779	Sequence 3779, App	367	14.2	71.0	393	18	US-10-425-115-130810	Sequence 130810, A
C 295	14.4	72.0	486	17	US-10-767-701-22567	Sequence 22567, App	368	14.2	71.0	393	18	US-10-425-115-130810	Sequence 130810, A
C 296	14.4	72.0	486	17	US-10-767-701-22567	Sequence 22567, App	369	14.2	71.0	419	18	US-10-425-115-127592	Sequence 127592, A
C 297	14.4	72.0	697	16	US-10-260-238-3781	Sequence 3781, App	370	14.2	71.0	419	18	US-10-425-115-127592	Sequence 127592, A
C 298	14.4	72.0	697	16	US-10-260-238-3781	Sequence 3781, App	371	14.2	71.0	427	18	US-10-425-115-173311	Sequence 173311, A
C 299	14.4	72.0	725	13	US-10-027-632-99538	Sequence 99538, A	372	14.2	71.0	427	18	US-10-425-115-173311	Sequence 173311, A
C 300	14.4	72.0	725	13	US-10-027-632-99538	Sequence 99538, A	373	14.2	71.0	456	9	US-09-738-626-941	Sequence 941, App
C 301	14.4	72.0	725	13	US-10-027-632-99538	Sequence 99538, A	374	14.2	71.0	456	9	US-09-738-626-941	Sequence 941, App
C 302	14.4	72.0	725	15	US-10-027-632-99538	Sequence 99538, A	375	14.2	71.0	459	18	US-10-425-115-55280	Sequence 55280, A
C 303	14.4	72.0	792	13	US-10-027-632-14096	Sequence 14096, A	376	14.2	71.0	459	18	US-10-425-115-55280	Sequence 55280, A
C 304	14.4	72.0	792	13	US-10-027-632-14096	Sequence 14096, A	377	14.2	71.0	464	18	US-10-425-115-82415	Sequence 82415, A

C 378	14.2	71.0	464	18	US-10-425-115-82415	Sequence 82415, A
C 379	14.2	71.0	476	13	US-10-037-632-50228	Sequence 50228, A
C 380	14.2	71.0	476	13	US-10-037-632-50228	Sequence 50228, A
C 381	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 382	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 383	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 384	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 385	14.2	71.0	476	13	US-10-027-632-50228	Sequence 50228, A
C 386	14.2	71.0	476	13	US-10-027-632-50228	Sequence 50228, A
C 387	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 388	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 389	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 390	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 391	14.2	71.0	485	18	US-10-425-115-161855	Sequence 161855, A
C 392	14.2	71.0	485	18	US-10-425-115-161855	Sequence 161855, A
C 393	14.2	71.0	531	16	US-10-425-114-7176	Sequence 7176, A
C 394	14.2	71.0	531	16	US-10-425-114-7176	Sequence 7176, A
C 395	14.2	71.0	531	16	US-10-425-115-108651	Sequence 108651, A
C 396	14.2	71.0	531	16	US-10-425-115-108651	Sequence 108651, A
C 397	14.2	71.0	563	16	US-10-260-238-3142	Sequence 3142, A
C 398	14.2	71.0	563	16	US-10-260-238-3142	Sequence 3142, A
C 399	14.2	71.0	580	16	US-10-260-238-4658	Sequence 4658, A
C 400	14.2	71.0	580	16	US-10-260-238-4658	Sequence 4658, A
C 401	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 402	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 403	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 404	14.2	71.0	583	13	US-10-027-632-113406	Sequence 113406, A
C 405	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 406	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 407	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 408	14.2	71.0	583	13	US-10-027-632-113406	Sequence 113406, A
C 409	14.2	71.0	589	16	US-10-260-238-3130	Sequence 3130, A
C 410	14.2	71.0	589	16	US-10-260-238-3130	Sequence 3130, A
C 411	14.2	71.0	590	18	US-10-357-930-59344	Sequence 59344, A
C 412	14.2	71.0	590	18	US-10-357-930-59344	Sequence 59344, A
C 413	14.2	71.0	591	15	US-10-006-285-309	Sequence 309, A
C 414	14.2	71.0	591	15	US-10-006-285-309	Sequence 309, A
C 415	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 416	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 417	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 418	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 419	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 420	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 421	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 422	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 423	14.2	71.0	600	15	US-10-027-632-204130	Sequence 204130, A
C 424	14.2	71.0	600	15	US-10-029-386-2368	Sequence 2368, A
C 425	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 426	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 427	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 428	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 429	14.2	71.0	611	13	US-10-027-632-82854	Sequence 82854, A
C 430	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 431	14.2	71.0	633	13	US-10-027-632-82855	Sequence 82855, A
C 432	14.2	71.0	633	13	US-10-027-632-82855	Sequence 82855, A
C 433	14.2	71.0	633	13	US-10-027-632-82855	Sequence 82855, A
C 434	14.2	71.0	633	13	US-10-027-632-82856	Sequence 82856, A
C 435	14.2	71.0	633	13	US-10-027-632-82856	Sequence 82856, A
C 436	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 437	14.2	71.0	633	13	US-10-027-632-82855	Sequence 82855, A
C 438	14.2	71.0	633	13	US-10-027-632-82855	Sequence 82855, A
C 439	14.2	71.0	633	13	US-10-027-632-82855	Sequence 82855, A
C 440	14.2	71.0	633	13	US-10-027-632-82856	Sequence 82856, A
C 441	14.2	71.0	636	16	US-10-260-238-3129	Sequence 3129, A
C 442	14.2	71.0	636	16	US-10-260-238-3129	Sequence 3129, A
C 443	14.2	71.0	636	16	US-10-653-047-6175	Sequence 6175, A
C 444	14.2	71.0	636	16	US-10-653-047-6175	Sequence 6175, A
C 445	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 446	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 447	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 448	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 449	14.2	71.0	657	15	US-10-282-122A-25895	Sequence 25895, A
C 450	14.2	71.0	657	15	US-10-282-122A-25895	Sequence 25895, A

## ALIGNMENTS

RESULT 1  
 US-10-447-136-162  
 ; Sequence 162, Application US/10447136  
 ; Publication No. US20040009948A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WRIGHT, Jim A.  
 ; APPLICANT: YOUNG, Arlindor  
 ; TITLE OF INVENTION: Antisense Sequences Directed Against R1 and  
 ; TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase  
 ; FILE REFERENCE: 032396-023  
 ; CURRENT APPLICATION NUMBER: US/10/447,136  
 ; PRIOR FILING DATE: 2003-05-29  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/249,247  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-11  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/023,040  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-02  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/039,959  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1997-03-07

;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 08/904,901  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-01  
;; NUMBER OF SEQ ID NOS: 220  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 162  
;; LENGTH: 20  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-447-136-162

Query Match 100.0%; Score 20; DB 16; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20  
DB 1 GGACATGCCGGGCGCATGTCC 20

## RESULT 2

US-10-447-136-162/c  
;; Sequence 162, Application US/10447136  
;; Publication No. US2004009948A1  
;; GENERAL INFORMATION:  
;; APPLICANT: WRIGHT, Jim A.  
;; APPLICANT: YOUNG, Apling H.  
;; TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and  
;; TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase  
;; FILE REFERENCE: 032396-023  
;; CURRENT APPLICATION NUMBER: US/10/447,136  
;; PRIOR FILING DATE: 2003-05-29  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/249,247  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-11  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/023,040  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-02  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/039,959  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-03-07  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 08/904,901  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-01  
;; NUMBER OF SEQ ID NOS: 220  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 162  
;; LENGTH: 20  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-447-136-162

Query Match 100.0%; Score 20; DB 16; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20  
DB 20 GGACATGCCGGGCGCATGTCC 1

## RESULT 3

US-10-023-318-19  
;; Sequence 19, Application US/10023318  
;; Publication No. US20030092015A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Larose, Anne-Marie  
;; APPLICANT: Rousseau, Pierre  
;; APPLICANT: Leblanc, Benoit  
;; APPLICANT: Camato, Rino  
;; TITLE OF INVENTION: Method for Screening and/or Identifying Factors that  
;; TITLE OF INVENTION: Bind to Nucleic Acids  
;; FILE REFERENCE: 9555.123USU1  
;; CURRENT APPLICATION NUMBER: US/10/023,318  
;; PRIOR FILING DATE: 2002-08-02  
;; PRIOR APPLICATION NUMBER: 2,327,561  
;; PRIOR FILING DATE: 2000-12-27  
;; NUMBER OF SEQ ID NOS: 41

;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 19  
;; LENGTH: 24  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: NABE-probes  
US-10-023-318-19

Query Match 100.0%; Score 20; DB 14; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20  
DB 5 GGACATGCCGGGCGCATGTCC 24

## RESULT 4

US-10-023-318-19/c  
;; Sequence 19, Application US/10023318  
;; Publication No. US20030092015A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Larose, Anne-Marie  
;; APPLICANT: Rousseau, Pierre  
;; APPLICANT: Leblanc, Benoit  
;; APPLICANT: Camato, Rino  
;; TITLE OF INVENTION: Method for Screening and/or Identifying Factors that  
;; TITLE OF INVENTION: Bind to Nucleic Acids  
;; FILE REFERENCE: 9555.123USU1  
;; CURRENT APPLICATION NUMBER: US/10/023,318  
;; PRIOR FILING DATE: 2002-08-02  
;; PRIOR APPLICATION NUMBER: 2,327,561  
;; PRIOR FILING DATE: 2000-12-27  
;; NUMBER OF SEQ ID NOS: 41  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 19  
;; LENGTH: 24  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: NABE-probes  
US-10-023-318-19

Query Match 100.0%; Score 20; DB 14; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20  
DB 24 GGACATGCCGGGCGCATGTCC 5

## RESULT 5

US-10-023-318-38  
;; Sequence 38, Application US/10023318  
;; Publication No. US20030092015A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Larose, Anne-Marie  
;; APPLICANT: Rousseau, Pierre  
;; APPLICANT: Leblanc, Benoit  
;; APPLICANT: Camato, Rino  
;; TITLE OF INVENTION: Method for Screening and/or Identifying Factors that  
;; TITLE OF INVENTION: Bind to Nucleic Acids  
;; FILE REFERENCE: 9555.123USU1  
;; CURRENT APPLICATION NUMBER: US/10/023,318  
;; PRIOR FILING DATE: 2002-08-02  
;; PRIOR APPLICATION NUMBER: 2,327,561  
;; PRIOR FILING DATE: 2000-12-27  
;; NUMBER OF SEQ ID NOS: 41  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 38  
;; LENGTH: 24

```
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Double
US-10-023-318-38
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGACATGCCCGGGCATGTCC 20
DB 5 GGACATGCCCGGGCATGTCC 24
```

```
RESULT 6
US-10-023-318-38/c
Sequence 38, Application US/10023318
Publication No. US20030092015A1
GENERAL INFORMATION:
APPLICANT: Larose, Anne-Marie
APPLICANT: Rousseau, Pierre
APPLICANT: Leblanc, Benoit
APPLICANT: Camato, Rino
TITLE OF INVENTION: Method for Screening and/or Identifying Factors that
TITLE OF INVENTION: Bind to Nucleic Acids
FILE REFERENCE: 9555.123USU1
CURRENT APPLICATION NUMBER: US/10/023,318
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: 2,327,561
PRIOR FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 38
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Double
US-10-023-318-38
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGACATGCCCGGGCATGTCC 20
DB 24 GGACATGCCCGGGCATGTCC 5
```

```
RESULT 7
US-10-339-161-14
Sequence 14, Application US/10339161
Publication No. US20030162211A1
GENERAL INFORMATION:
APPLICANT: Remacle, Jose
APPLICANT: Renard, Patricia
APPLICANT: Art, Muriel
TITLE OF INVENTION: METHOD AND KIT FOR THE DETERMINATION OF
FILE REFERENCE: VANM212.001CPI
CURRENT APPLICATION NUMBER: US/10/339,161
CURRENT FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: US 09/816,763
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: EP 00870057.7
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 19
```

```
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Transcription factor p53
US-10-339-161-14
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 15; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGACATGCCCGGGCATGTTC 19
DB 1 GGACATGCCCGGGCATGTTC 19
```

```
RESULT 8
US-10-339-161-14/c
Sequence 14, Application US/10339161
Publication No. US20030162211A1
GENERAL INFORMATION:
APPLICANT: Remacle, Jose
APPLICANT: Renard, Patricia
APPLICANT: Art, Muriel
TITLE OF INVENTION: METHOD AND KIT FOR THE DETERMINATION OF
TITLE OF INVENTION: CELLULAR ACTIVATION PROFILES
FILE REFERENCE: VANM212.001CPI
CURRENT APPLICATION NUMBER: US/10/339,161
CURRENT FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: US 09/816,763
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: EP 00870057.7
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 19
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Transcription factor p53
US-10-339-161-14
```

```
Query Match
Best Local Similarity 95.0%; Score 19; DB 15; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 GACATGCCCGGGCATGTCC 20
DB 19 GACATGCCCGGGCATGTCC 1
```

```
RESULT 9
US-10-767-701-2482
Sequence 2482, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 2482
LENGTH: 729
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS41949_1
US-10-767-701-2482
```

Query Match 87.0%; Score 17.4; DB 17; Length 729;  
Best Local Similarity 94.7%; Pred. No. 34;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGACATGCCCGGGCATGTC 19  
DB 247 GGACACGCCCGGGCATGTC 265

RESULT 10  
US-10-767-701-2482/c  
; Sequence 2482, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 2482  
; LENGTH: 729  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS41949\_1  
US-10-767-701-2482

Query Match 87.0%; Score 17.4; DB 17; Length 729;  
Best Local Similarity 94.7%; Pred. No. 34;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACATGCCCGGGCATGTC 20  
DB 265 GACATGCCCGGGCATGTC 247

RESULT 11  
US-10-425-114-30682  
; Sequence 30682, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 30682  
; LENGTH: 3325  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73063H08\_F11  
US-10-425-114-30682

Query Match 87.0%; Score 17.4; DB 16; Length 3325;  
Best Local Similarity 94.7%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGACATGCCCGGGCATGTC 19  
DB 739 GGACACGCCCGGGCATGTC 757

RESULT 12  
US-10-425-114-30682/c  
; Sequence 30682, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 30682  
; LENGTH: 3325  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73063H08\_F11  
US-10-425-114-30682

Query Match 87.0%; Score 17.4; DB 16; Length 3325;  
Best Local Similarity 94.7%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACATGCCCGGGCATGTC 20  
DB 757 GACATGCCCGGGCATGTC 739

RESULT 13  
US-10-425-115-4271  
; Sequence 4271, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 4271  
; LENGTH: 3458  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_103895C.1  
US-10-425-115-4271

Query Match 87.0%; Score 17.4; DB 18; Length 3458;  
Best Local Similarity 94.7%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGACATGCCCGGGCATGTC 19  
DB 739 GGACACGCCCGGGCATGTC 757

RESULT 14  
US-10-425-115-4271/c  
; Sequence 4271, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 4271  
LENGTH: 3458  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MFT4577\_103895C.1  
US-10-425-115-4271

Query Match 87.0%; Score 17.4; DB 16; Length 3458;  
Best Local Similarity 94.7%; Pred. No. 28;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGCGCATGTC 20  
DB 757 GACATGCCCGCGCGCATGTC 739

RESULT 15  
US-10-425-114-2765  
Sequence 2765, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 2765  
LENGTH: 3569  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700223564\_FLI  
US-10-425-114-2765

Query Match 87.0%; Score 17.4; DB 16; Length 3569;  
Best Local Similarity 94.7%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGCGCATGTC 19  
DB 733 GGACATGCCCGCGCATGTC 751

RESULT 16  
US-10-425-114-2765/C  
Sequence 2765, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 2765  
LENGTH: 3569  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700223564\_FLI  
US-10-425-114-2765

FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 2765  
LENGTH: 3569  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700223564\_FLI  
US-10-425-114-2765

Query Match 87.0%; Score 17.4; DB 16; Length 3569;  
Best Local Similarity 94.7%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGCGCATGTC 20  
DB 751 GACATGCCCGCGCGCATGTC 733

RESULT 17  
US-10-425-115-4270  
Sequence 4270, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 4270  
LENGTH: 3609  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MFT4577\_103894C.1  
US-10-425-115-4270

Query Match 87.0%; Score 17.4; DB 18; Length 3609;  
Best Local Similarity 94.7%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGCGCATGTC 19  
DB 755 GGACATGCCCGCGCATGTC 773

RESULT 18  
US-10-425-115-4270/C  
Sequence 4270, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 4270  
LENGTH: 3609  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:



OTHER INFORMATION: Clone ID: MNT4577\_103894C.1  
US-10-425-115-4270

Query Match 84.0%; Score 16.8; DB 10; Length 30;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGACATGCCCGGGCATGTCC 20  
DB 773 GGACATGCCCGGGCATGTCC 755

RESULT 19  
US-09-829-922-31

Sequence 31, Application US/09829922  
Publication No. US20030171537A1

GENERAL INFORMATION:

APPLICANT: Halazoneis, Thanos

APPLICANT: Hartwig, Wolfgang

TITLE OF INVENTION: Peptides and peptidomimetics with

TITLE OF INVENTION: structural similarity to human p53 that activate p53

TITLE OF INVENTION: function

FILE REFERENCE: 2973.19998

CURRENT APPLICATION NUMBER: US/09/829,922

CURRENT FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 08/894,327

PRIOR FILING DATE: 1997-12-04

PRIOR APPLICATION NUMBER: pctus96/01535

PRIOR FILING DATE: 1996-02-16

PRIOR APPLICATION NUMBER: 08/392,542

PRIOR FILING DATE: 1995-02-16

NUMBER OF SEQ ID NOS: 35

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 31

LENGTH: 30

TYPE: DNA

ORGANISM: Homo sapiens

US-09-829-922-31

Query Match 84.0%; Score 16.8; DB 10; Length 30;  
Best Local Similarity 90.0%; Pred. No. 99;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
DB 3 GGACATGTCCGGGCATGTCC 22

RESULT 20  
US-09-829-922-31/c

Sequence 31, Application US/09829922  
Publication No. US20030171537A1

GENERAL INFORMATION:

APPLICANT: Halazoneis, Thanos

APPLICANT: Hartwig, Wolfgang

TITLE OF INVENTION: Peptides and peptidomimetics with

TITLE OF INVENTION: structural similarity to human p53 that activate p53

TITLE OF INVENTION: function

FILE REFERENCE: 2973.19998

CURRENT APPLICATION NUMBER: US/09/829,922

CURRENT FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 08/894,327

PRIOR FILING DATE: 1997-12-04

PRIOR APPLICATION NUMBER: pctus96/01535

PRIOR FILING DATE: 1996-02-16

PRIOR APPLICATION NUMBER: 08/392,542

PRIOR FILING DATE: 1995-02-16

NUMBER OF SEQ ID NOS: 35

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 31

LENGTH: 30

TYPE: DNA

ORGANISM: Homo sapiens

US-09-829-922-31

Query Match 84.0%; Score 16.8; DB 10; Length 30;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
DB 22 GGACATGCCCGGGCATGTCC 3

RESULT 21  
US-10-160-290-31

Sequence 31, Application US/10160290  
Publication No. US20030124557A1

GENERAL INFORMATION:

APPLICANT: Halazoneis, Thanos

APPLICANT: Hartwig, Wolfgang

TITLE OF INVENTION: Peptides and peptidomimetics with

TITLE OF INVENTION: structural similarity to Human p53 That Activate

TITLE OF INVENTION: function

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner, Birch, McKie & Beckett

STREET: 1001 G Street, N.W.

CITY: Washington, D.C.

STATE: District of Columbia

COUNTRY: U.S.

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/160,290

FILING DATE: 04-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/685,027

FILING DATE: 10-Oct-2000

APPLICATION NUMBER: 08/392,542

FILING DATE: <Unknown>

AUTHOR/AGENT INFORMATION:

NAME: Boborske, Laurence H.

REGISTRATION NUMBER: 34,698

REFERENCE/DOCKET NUMBER: 0486,48439

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202 508-9100

TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-10-160-290-31

Query Match 84.0%; Score 16.8; DB 15; Length 30;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20  
DB 3 GGACATGTCCGGGCATGTCC 22

RESULT 22

US-10-160-290-31/c  
; Sequence 31, Application US/10160290  
; Publication No. US20030124557A1  
; GENERAL INFORMATION:  
; APPLICANT: Halazonetis, Thanos  
; TITLE OF INVENTION: Peptides nad Peptidomimetics with  
; Structural Similarity to Human p53 That Activate  
; p53  
; Function  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie & Beckett  
; STREET: 1001 G Street, N.W.  
; CITY: Washington, D.C.  
; STATE: District of Columbia  
; COUNTRY: U.S.  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/160,290  
; FILING DATE: 04-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/685,027  
; FILING DATE: 10-Oct-2000  
; APPLICATION NUMBER: 08/392,542  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Posorske, Laurence H.  
; REGISTRATION NUMBER: 34,698  
; REFERENCE/DOCKET NUMBER: 0486.48439  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 508-9100  
; TELEFAX: 202 508-9299  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 30 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-10-160-290-31  
Query Match 84.0%; Score 16.8; DB 15; Length 30;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Cy 1 GGACATGCCCGGCGCATGTCC 20  
Db 22 GGACATGCCCGGCGCATGTCC 3  
RESULT 23  
US-10-052-482-61  
; Sequence 61, Application US/10052482  
; Publication No. US20040072264A1  
; GENERAL INFORMATION:  
; APPLICANT: Engelhard, Eric  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER  
; FILE REFERENCE: A-71087/RMS/DCF  
; CURRENT APPLICATION NUMBER: US/10/052,482  
; CURRENT FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 61  
; LENGTH: 21565  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1217)..(1236)  
; OTHER INFORMATION: "n" at positions 1217 to 1236 can be any base  
; NAME/KEY: misc\_feature  
; LOCATION: (3121)..(3140)  
; OTHER INFORMATION: "n" at positions 3121 to 3140 can be any base  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (5706)..(5725)  
; OTHER INFORMATION: "n" at positions 5706 to 5725 can be any base  
US-10-052-482-61  
Query Match 84.0%; Score 16.8; DB 16; Length 21565;  
Best Local Similarity 90.0%; Pred. No. 45;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Cy 1 GGACATGCCCGGCGCATGTCC 20  
Db 9378 GGCGTTCGCCGGCGCATGTCC 9378  
RESULT 24  
US-10-052-482-61/c  
; Sequence 61, Application US/10052482  
; Publication No. US20040072264A1  
; GENERAL INFORMATION:  
; APPLICANT: Engelhard, Eric  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER  
; FILE REFERENCE: A-71087/RMS/DCF  
; CURRENT APPLICATION NUMBER: US/10/052,482  
; CURRENT FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 61  
; LENGTH: 21565  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1217)..(1236)  
; OTHER INFORMATION: "n" at positions 1217 to 1236 can be any base  
; NAME/KEY: misc\_feature  
; LOCATION: (3121)..(3140)  
; OTHER INFORMATION: "n" at positions 3121 to 3140 can be any base  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (5706)..(5725)  
; OTHER INFORMATION: "n" at positions 5706 to 5725 can be any base  
US-10-052-482-61  
Query Match 84.0%; Score 16.8; DB 16; Length 21565;  
Best Local Similarity 90.0%; Pred. No. 45;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Cy 1 GGACATGCCCGGCGCATGTCC 20  
Db 9397 GGCGTTCGCCGGCGCATGTCC 9378

```
RESULT 25
US-10-367-094-1
; Sequence 1, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 46030
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(46030)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-1

Query Match      84.0%; Score 16.8; DB 17; Length 46030;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTCC 20
Db      6907 GGGCTGTCCCGGCGCATGTCC 6926

RESULT 26
US-10-367-094-1/c
; Sequence 1, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 46030
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(46030)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-1

Query Match      84.0%; Score 16.8; DB 17; Length 46030;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTCC 20
Db      6926 GGACATGCCCGGCGCATGTCC 6907
```

```
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 187
; LENGTH: 60327
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1943)..(1968)
; OTHER INFORMATION: "n" at positions 1949 to 1968 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (8743)..(8762)
; OTHER INFORMATION: "n" at positions 8743 to 8762 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (14521)..(14618)
; OTHER INFORMATION: "n" at positions 14521 to 14618 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (18936)..(18955)
; OTHER INFORMATION: "n" at positions 18936 to 18955 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (25819)..(26117)
; OTHER INFORMATION: "n" at positions 25819 to 26117 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (38774)..(39110)
; OTHER INFORMATION: "n" at positions 38774 to 39110 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (48067)..(48086)
; OTHER INFORMATION: "n" at positions 48067 to 48086 can be any base
US-10-052-482-187

Query Match      84.0%; Score 16.8; DB 16; Length 60327;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTCC 20
Db      3740 GGACATGCCCGGCGCATGTCC 3759

RESULT 28
US-10-052-482-187/c
; Sequence 187, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 187
; LENGTH: 60327
; TYPE: DNA
```

```

; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1945)..(1958)
; OTHER INFORMATION: "n" at positions 1949 to 1968 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8743)..(8762)
; OTHER INFORMATION: "n" at positions 8743 to 8762 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14521)..(14618)
; OTHER INFORMATION: "n" at positions 14521 to 14618 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18936)..(18955)
; OTHER INFORMATION: "n" at positions 18936 to 18955 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (25819)..(26117)
; OTHER INFORMATION: "n" at positions 25819 to 26117 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (38774)..(39110)
; OTHER INFORMATION: "n" at positions 38774 to 39110 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (48067)..(48086)
; OTHER INFORMATION: "n" at positions 48067 to 48086 can be any base
; US-10-052-482-187
```

```

Query Match      84.0%; Score 16.8; DB 16; Length 60327;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      3759 GGGCCTGCCCGGGCATGTCC 3740
```

```

RESULT 29
US-10-034-650-31
; Sequence 31, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 68233
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-034-650-31
```

```

Query Match      84.0%; Score 16.8; DB 15; Length 68233;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      14433 GGACATGCCCGGGCATGTCC 14452
```

```

RESULT 30
US-10-034-650-31/c
; Sequence 31, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 68233
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-034-650-31
```

```

Query Match      84.0%; Score 16.8; DB 15; Length 68233;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      14452 GGGCCTGCCCGGGCATGTCC 14433
```

```

RESULT 31
US-10-087-192-709
; Sequence 709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 709
; LENGTH: 77941
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..((77941))
; OTHER INFORMATION: n = A,T,C or G
; US-10-087-192-709
```

```

Query Match      84.0%; Score 16.8; DB 13; Length 77941;
Best Local Similarity 90.0%; Pred. No. 39;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      65117 GGGCCTGCCCGGGCATGTCC 65136
```

```

RESULT 32
US-10-087-192-709/c
; Sequence 709, Application US/10087192
```

```
Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 709
; LENGTH: 77941
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(77941)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-709

Query Match      84.0%; Score 16.8; DB 13; Length 77941;
Best Local Similarity 90.0%; Pred. No. 39;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 GGACATGCCCGGCGCATGTCC 20
Db      65136 GGACATGCCCGGCGCATGTCC 65117

RESULT 33
US-10-322-281-33
; Sequence 33, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 145068
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(145068)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-33

Query Match      84.0%; Score 16.8; DB 17; Length 145068;
Best Local Similarity 90.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 GGACATGCCCGGCGCATGTCC 20
Db      95566 GGACATGCCCGGCGCATGTCC 95565
```

```
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 145068
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(145068)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-33

Query Match      84.0%; Score 16.8; DB 17; Length 145068;
Best Local Similarity 90.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 GGACATGCCCGGCGCATGTCC 20
Db      95565 GGACATGCCCGGCGCATGTCC 95566

RESULT 35
US-09-835-976B-77
; Sequence 77, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 77
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-835-976B-77

Query Match      82.0%; Score 16.4; DB 10; Length 400;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      3 ACATGCCCGGCGCATGTCC 20
Db      351 ACATGCCCGGCGCATGTCC 368

RESULT 36
US-09-835-976B-77/c
; Sequence 77, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
```

SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 77  
LENGTH: 400  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-835-976B-77

Query Match 82.0%; Score 16.4; DB 10; Length 400;  
Best Local Similarity 94.4%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18  
DB 368 GGACATGCCCGGCATGT 351

RESULT 37  
US-09-835-976B-111  
Sequence 111, Application US/09835976B  
Publication No. US20030027983A1  
GENERAL INFORMATION:  
APPLICANT: Mount, David B.  
APPLICANT: Delplre, Eric  
APPLICANT: Gamba, Gerardo  
APPLICANT: Alfred L. George, Jr.  
TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC  
TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME  
FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2  
CURRENT FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 131  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 111  
LENGTH: 3422  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY:  
LOCATION:  
OTHER INFORMATION: n=a, c, g, or t  
US-09-835-976B-111

Query Match 82.0%; Score 16.4; DB 10; Length 3422;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACATGCCCGGCATGTCC 20  
DB 3051 ACATGCCCGGCATGTCC 3068

RESULT 38  
US-09-835-976B-111/c  
Sequence 111, Application US/09835976B  
Publication No. US20030027983A1  
GENERAL INFORMATION:  
APPLICANT: Mount, David B.  
APPLICANT: Delplre, Eric  
APPLICANT: Gamba, Gerardo  
APPLICANT: Alfred L. George, Jr.  
TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC  
TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME  
FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2  
CURRENT FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 131  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 111  
LENGTH: 3422  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY:  
LOCATION:  
OTHER INFORMATION: n=a, c, g, or t  
US-09-835-976B-111

Query Match 82.0%; Score 16.4; DB 10; Length 3422;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18  
DB 3068 GGACATGCCCGGCATGT 3051

RESULT 39  
US-10-322-696-40  
Sequence 40, Application US/10322696  
Publication No. US20040166490A1  
GENERAL INFORMATION:  
APPLICANT: Morrie, David W.  
APPLICANT: Malandro, Marc  
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER  
FILE REFERENCE: 529452001200  
CURRENT APPLICATION NUMBER: US/10/322,696  
CURRENT FILING DATE: 2003-10-17  
NUMBER OF SEQ ID NOS: 186  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 40  
LENGTH: 32351  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(32351)  
OTHER INFORMATION: n = A,T,C or G  
US-10-322-696-40

Query Match 82.0%; Score 16.4; DB 17; Length 32351;  
Best Local Similarity 94.4%; Pred. No. 69;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACATGCCCGGCATGTCC 20  
DB 322 ACATGCCCGGCATGTCC 339

RESULT 40  
US-10-322-696-40/c  
Sequence 40, Application US/10322696  
Publication No. US20040166490A1  
GENERAL INFORMATION:  
APPLICANT: Morrie, David W.  
APPLICANT: Malandro, Marc  
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER  
FILE REFERENCE: 529452001200  
CURRENT APPLICATION NUMBER: US/10/322,696  
CURRENT FILING DATE: 2003-10-17  
NUMBER OF SEQ ID NOS: 186  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 40  
LENGTH: 32351  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(32351)  
OTHER INFORMATION: n = A,T,C or G  
US-10-322-696-40

Query Match 82.0%; Score 16.4; DB 17; Length 32351;  
Best Local Similarity 94.4%; Pred. No. 69;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18  
|||  
Db 339 GGACATGCCCGGCATGT 322

RESULT 41  
US-10-087-192-988  
; Sequence 988, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 988  
; LENGTH: 203264  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-087-192-988

Query Match 82.0%; Score 16.4; DB 13; Length 203264;  
Best Local Similarity 94.4%; Pred. No. 55;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACATGCCCGGCATGTCC 20  
|||  
Db 42121 ACATGCCCGGCATGTCC 42138

RESULT 42  
US-10-087-192-988/c  
; Sequence 988, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 988  
; LENGTH: 203264  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-087-192-988

Query Match 82.0%; Score 16.4; DB 13; Length 203264;  
Best Local Similarity 94.4%; Pred. No. 55;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18  
|||  
Db 42138 GGACATGCCCGGCATGT 42121

RESULT 43

US-09-294-093B-1681  
; Sequence 1681, Application US/09294093B  
; Patent No. US20010051335A1  
; GENERAL INFORMATION:  
; APPLICANT: Ito, Laura, Y.  
; APPLICANT: Sherman, Bradley, K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
; FILE REFERENCE: PL-0009 US  
; CURRENT APPLICATION NUMBER: US/09/294,093B  
; CURRENT FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/082,567  
; PRIOR FILING DATE: April 21, 1998  
; NUMBER OF SEQ ID NOS: 6207  
; SOFTWARE: PERL Program  
; SEQ ID NO 1681  
; LENGTH: 274  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700344792H1  
US-09-294-093B-1681

Query Match 79.0%; Score 15.8; DB 9; Length 274;  
Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTTC 19  
|||  
Db 189 GGACATGCCCGGCATGTTC 207

RESULT 44  
US-09-294-093B-1681/c  
; Sequence 1681, Application US/09294093B  
; Patent No. US20010051335A1  
; GENERAL INFORMATION:  
; APPLICANT: Ito, Laura, Y.  
; APPLICANT: Sherman, Bradley, K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
; FILE REFERENCE: PL-0009 US  
; CURRENT APPLICATION NUMBER: US/09/294,093B  
; CURRENT FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/082,567  
; PRIOR FILING DATE: April 21, 1998  
; NUMBER OF SEQ ID NOS: 6207  
; SOFTWARE: PERL Program  
; SEQ ID NO 1681  
; LENGTH: 274  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700344792H1  
US-09-294-093B-1681

Query Match 79.0%; Score 15.8; DB 9; Length 274;  
Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCATGTTC 20  
|||  
Db 207 GACATGCCCGGCATGTTC 189

RESULT 45  
US-09-991-936-580  
; Sequence 580, Application US/09991936  
; Publication No. US20030073827A1  
; GENERAL INFORMATION:  
; APPLICANT: Brandt, Kevin S.

```

; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 580
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-580

```

```

Query Match      79.0%; Score 15.8; DB 10; Length 277;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GGACATGCCCGGCGCATGTC 19
DB 129 GGACATTCCTCGGCATGTC 147

```

```

RESULT 46
US-09-991-936-580/c
; Sequence 580, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 580
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-580

```

```

Query Match      79.0%; Score 15.8; DB 10; Length 277;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 GACATGCCCGGCGCATGTC 20
DB 147 GACATGCCCGGCGCATGTC 129

```

```

RESULT 47
US-09-294-093B-5575
; Sequence 5575, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL

```

```

; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5575
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700357426H1
; LOCATION: 69
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-5575

```

```

Query Match      79.0%; Score 15.8; DB 9; Length 294;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GGACATGCCCGGCGCATGTC 19
DB 78 GGACATTCCTCGGCATGTC 96

```

```

RESULT 48
US-09-294-093B-5575/c
; Sequence 5575, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5575
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700357426H1
; LOCATION: 69
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-5575

```

```

Query Match      79.0%; Score 15.8; DB 9; Length 294;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 GACATGCCCGGCGCATGTC 20
DB 96 GACATGCCCGGCGCATGTC 78

```

```

RESULT 49
US-09-783-590-4831
; Sequence 4831, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong

```



```

; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIORITY FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4831
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (218)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (252)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (288)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (353)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (417)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (447)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (466)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (468)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (474)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-4831

Query Match          79.0%: Score 15.8; DB 9; Length 482;
Best Local Similarity 89.5%: Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTC 19
Db      148 GGACATGCCCGGCGCATGCC 166
```

```

; Patent No.: US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIORITY FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4831
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (218)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (252)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (288)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (353)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (417)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (447)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (466)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (468)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (474)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-4831

Query Match          79.0%: Score 15.8; DB 9; Length 482;
Best Local Similarity 89.5%: Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GACATGCCCGGCGCATGTC 20
Db      166 GGCATTCGCGGCGCATGTC 148
```

Sun Jan 16 16:27:32 2005

us-09-578-453-2.rnpb

Page 18

Search completed: January 14, 2005, 19:48:25  
Job time : 381.421 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:35:09 ; Search time 2632.63 Seconds  
(without alignments)  
276.831 Million cell updates/sec

Title: US-09-578-453-2

Perfect score: 20

Sequence: 1 GGACATGCCCGCGCATGTC 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	683	9	CR249440 Forward s
2	18.4	92.0	683	9	CR249440 Forward s
3	17.4	87.0	409	8	BH781927 f2mb001f0
4	17.4	87.0	409	8	BH781927 f2mb001f0
5	17.4	87.0	482	4	BG556896 EM1_40_AO
6	17.4	87.0	482	4	BG556896 EM1_40_AO
7	17.4	87.0	557	4	BI098283 IP1_30_D1
8	17.4	87.0	557	4	BI098283 IP1_30_D1
9	17.4	87.0	757	9	CG270390 OG3D059TV
10	17.4	87.0	757	9	CG270390 OG3D059TV
11	17.4	87.0	843	9	CG290290 OG2BC47TH
12	17.4	87.0	843	9	CG290290 OG2BC47TH
13	17.4	87.0	870	9	CG236275 OG1CU67TH
14	17.4	87.0	870	9	CG236275 OG1CU67TH
15	17.4	87.0	876	8	BZ364970 PUD6823TD
16	17.4	87.0	876	8	BZ364970 PUD6823TD
17	17.4	87.0	877	9	CL637471 CH243-SH5
18	17.4	87.0	877	9	CL637471 CH243-SH5
19	17.4	87.0	930	9	CG336287 OG1CU67TV
20	17.4	87.0	930	9	CG336287 OG1CU67TV
21	17.4	87.0	977	9	CG290302 OG2BC47TV
22	17.4	87.0	977	9	CG290302 OG2BC47TV
23	16.8	84.0	179	9	CL640722 M003D05 G
24	16.8	84.0	179	9	CL640722 M003D05 G

25	16.8	84.0	221	8	BZ691000 M003D05 G
26	16.8	84.0	221	8	BZ691000 M003D05 G
27	16.8	84.0	227	1	AV322349 AV322349
28	16.8	84.0	227	1	AV322349 AV322349
29	16.8	84.0	269	8	AZ377623 IM0132L01
30	16.8	84.0	269	8	AZ377623 IM0132L01
31	16.8	84.0	274	8	AZ505789 IM0346R12
32	16.8	84.0	274	8	AZ505789 IM0346R12
33	16.8	84.0	280	5	BP090312 BP090312
34	16.8	84.0	280	5	BP090312 BP090312
35	16.8	84.0	282	2	BB093079 BB093079
36	16.8	84.0	282	2	BB093079 BB093079
37	16.8	84.0	283	2	BB157894 BB157894
38	16.8	84.0	283	2	BB157894 BB157894
39	16.8	84.0	283	2	BB755436 BB755436
40	16.8	84.0	283	2	BB755436 BB755436
41	16.8	84.0	287	5	BY469224 BY469224
42	16.8	84.0	287	5	BY469224 BY469224
43	16.8	84.0	288	7	CK337629 CK337629
44	16.8	84.0	288	7	CK337629 CK337629
45	16.8	84.0	288	7	CK340780 CK340780
46	16.8	84.0	288	7	CK340780 CK340780
47	16.8	84.0	290	2	BB060455 BB060455
48	16.8	84.0	290	2	BB060455 BB060455
49	16.8	84.0	293	1	AA289954 AA289954
50	16.8	84.0	293	1	AA289954 AA289954
51	16.8	84.0	303	6	CD541539 CD541539
52	16.8	84.0	303	6	CD541539 CD541539
53	16.8	84.0	306	6	CD539301 CD539301
54	16.8	84.0	306	6	CD539301 CD539301
55	16.8	84.0	307	2	AM213731 AM213731
56	16.8	84.0	307	2	AM213731 AM213731
57	16.8	84.0	311	2	BB154112 BB154112
58	16.8	84.0	311	2	BB154112 BB154112
59	16.8	84.0	319	2	BB132471 BB132471
60	16.8	84.0	319	2	BB132471 BB132471
61	16.8	84.0	321	1	AA408122 AA408122
62	16.8	84.0	321	1	AA408122 AA408122
63	16.8	84.0	324	7	R61912 R61912
64	16.8	84.0	324	7	R61912 R61912
65	16.8	84.0	327	2	BF452129 BF452129
66	16.8	84.0	327	2	BF452129 BF452129
67	16.8	84.0	333	1	AA445841 AA445841
68	16.8	84.0	333	1	AA445841 AA445841
69	16.8	84.0	333	7	CK337615 CK337615
70	16.8	84.0	333	7	CK337615 CK337615
71	16.8	84.0	338	7	CK337751 CK337751
72	16.8	84.0	338	7	CK337751 CK337751
73	16.8	84.0	346	1	AA617393 AA617393
74	16.8	84.0	346	1	AA617393 AA617393
75	16.8	84.0	347	2	BB670724 BB670724
76	16.8	84.0	347	2	BB670724 BB670724
77	16.8	84.0	351	8	AQ992743 RPCI-23-2
78	16.8	84.0	351	8	AQ992743 RPCI-23-2
79	16.8	84.0	358	1	AV117495 AV117495
80	16.8	84.0	358	1	AV117495 AV117495
81	16.8	84.0	361	2	BB802219 BB802219
82	16.8	84.0	361	2	BB802219 BB802219
83	16.8	84.0	366	2	BB809747 BB809747
84	16.8	84.0	366	2	BB809747 BB809747
85	16.8	84.0	369	1	AA427020 AA427020
86	16.8	84.0	369	1	AA427020 AA427020
87	16.8	84.0	376	8	AZ483669 AZ483669
88	16.8	84.0	376	8	AZ483669 AZ483669
89	16.8	84.0	377	1	AI550171 AI550171
90	16.8	84.0	377	1	AI550171 AI550171
91	16.8	84.0	379	2	AM323086 AM323086
92	16.8	84.0	379	2	AM323086 AM323086
93	16.8	84.0	382	2	AM741148 AM741148
94	16.8	84.0	382	2	AM741148 AM741148
95	16.8	84.0	382	2	BB793888 BB793888
96	16.8	84.0	382	2	BB793888 BB793888
97	16.8	84.0	390	2	BB800634 BB800634

C 98	16.8	84.0	390 2	BB80634	BB80634	171	16.8	84.0	459 1	A1159253
C 99	16.8	84.0	390 7	CK341151	CK341151	C 172	16.8	84.0	459 1	A1159253
C 100	16.8	84.0	390 7	CK341151	CK341151	C 173	16.8	84.0	462 6	C87458
C 101	16.8	84.0	391 7	CP947326	CP947326	C 174	16.8	84.0	462 6	C87458
C 102	16.8	84.0	391 7	CP947326	CP947326	C 175	16.8	84.0	462 6	C87458
C 103	16.8	84.0	391 8	A2240326	A2240326	C 176	16.8	84.0	462 9	CL631989
C 104	16.8	84.0	391 8	A2240326	A2240326	C 177	16.8	84.0	465 2	BE369064
C 105	16.8	84.0	393 7	CK340826	CK340826	C 178	16.8	84.0	465 2	BE369064
C 106	16.8	84.0	393 7	CK340826	CK340826	C 179	16.8	84.0	465 7	CK340771
C 107	16.8	84.0	394 1	AA125901	AA125901	C 180	16.8	84.0	465 7	CK340771
C 108	16.8	84.0	394 1	AA125901	AA125901	C 181	16.8	84.0	466 8	AZ646335
C 109	16.8	84.0	394 7	CK342267	CK342267	C 182	16.8	84.0	466 8	AZ646335
C 110	16.8	84.0	394 7	CK342267	CK342267	C 183	16.8	84.0	467 1	A1592934
C 111	16.8	84.0	397 8	A2788367	A2788367	C 184	16.8	84.0	467 1	A1592934
C 112	16.8	84.0	397 8	A2788367	A2788367	C 185	16.8	84.0	467 6	CA667592
C 113	16.8	84.0	398 6	CD548288	CD548288	C 186	16.8	84.0	467 6	CA667592
C 114	16.8	84.0	399 1	AI145059	AI145059	C 187	16.8	84.0	468 8	AZ031229
C 115	16.8	84.0	399 1	AI145059	AI145059	C 188	16.8	84.0	468 8	AZ031229
C 116	16.8	84.0	399 1	AI145059	AI145059	C 189	16.8	84.0	470 2	BB700662
C 117	16.8	84.0	402 2	BB802309	BB802309	C 190	16.8	84.0	470 2	BB700662
C 118	16.8	84.0	402 2	BB802309	BB802309	C 191	16.8	84.0	470 2	BB700662
C 119	16.8	84.0	406 2	AA741532	AA741532	C 192	16.8	84.0	470 2	BB700662
C 120	16.8	84.0	406 2	AA741532	AA741532	C 193	16.8	84.0	473 5	BO886865
C 121	16.8	84.0	406 2	BB742521	BB742521	C 194	16.8	84.0	473 5	BO886865
C 122	16.8	84.0	406 2	BB742521	BB742521	C 195	16.8	84.0	477 8	AZ264222
C 123	16.8	84.0	407 5	BY546952	BY546952	C 196	16.8	84.0	477 8	AZ264222
C 124	16.8	84.0	408 2	AA206072	AA206072	C 197	16.8	84.0	487 2	BB760838
C 125	16.8	84.0	408 2	AA206072	AA206072	C 198	16.8	84.0	487 2	BB760838
C 126	16.8	84.0	408 5	BB264927	BB264927	C 199	16.8	84.0	489 8	AZ585443
C 127	16.8	84.0	408 5	BB264927	BB264927	C 200	16.8	84.0	489 8	AZ585443
C 128	16.8	84.0	408 7	W77307	W77307	C 201	16.8	84.0	492 2	AA470831
C 129	16.8	84.0	408 7	W77307	W77307	C 202	16.8	84.0	492 2	AA470831
C 130	16.8	84.0	412 8	BH031966	BH031966	C 203	16.8	84.0	493 8	BH055085
C 131	16.8	84.0	412 8	BH031966	BH031966	C 204	16.8	84.0	493 8	BH055085
C 132	16.8	84.0	412 8	BH031966	BH031966	C 205	16.8	84.0	497 8	AZ989198
C 133	16.8	84.0	415 2	BB788289	BB788289	C 206	16.8	84.0	497 8	AZ989198
C 134	16.8	84.0	415 2	BB788289	BB788289	C 207	16.8	84.0	500 7	CK361147
C 135	16.8	84.0	419 1	AA967383	AA967383	C 208	16.8	84.0	500 7	CK361147
C 136	16.8	84.0	419 1	AA967383	AA967383	C 209	16.8	84.0	501 8	AZ646304
C 137	16.8	84.0	422 2	BB671022	BB671022	C 210	16.8	84.0	501 8	AZ646304
C 138	16.8	84.0	422 2	BB671022	BB671022	C 211	16.8	84.0	503 8	AZ080197
C 139	16.8	84.0	423 9	CNS07EUL	CNS07EUL	C 212	16.8	84.0	503 8	AZ080197
C 140	16.8	84.0	423 9	CNS07EUL	CNS07EUL	C 213	16.8	84.0	512 2	BB755827
C 141	16.8	84.0	427 2	BB790475	BB790475	C 214	16.8	84.0	512 2	BB755827
C 142	16.8	84.0	427 2	BB790475	BB790475	C 215	16.8	84.0	514 2	BB700756
C 143	16.8	84.0	430 1	AU017612	AU017612	C 216	16.8	84.0	514 2	BB700756
C 144	16.8	84.0	430 1	AU017612	AU017612	C 217	16.8	84.0	516 8	AZ587847
C 145	16.8	84.0	431 2	BB727075	BB727075	C 218	16.8	84.0	516 8	AZ587847
C 146	16.8	84.0	431 2	BB727075	BB727075	C 219	16.8	84.0	520 8	AZ031430
C 147	16.8	84.0	431 2	BB727075	BB727075	C 220	16.8	84.0	520 8	AZ031430
C 148	16.8	84.0	431 2	BB727075	BB727075	C 221	16.8	84.0	521 2	AA060682
C 149	16.8	84.0	432 2	BB670907	BB670907	C 222	16.8	84.0	521 2	AA060682
C 150	16.8	84.0	432 2	BB670907	BB670907	C 223	16.8	84.0	521 2	BB755859
C 151	16.8	84.0	435 2	AA763751	AA763751	C 224	16.8	84.0	521 2	BB755859
C 152	16.8	84.0	435 2	AA763751	AA763751	C 225	16.8	84.0	522 2	BB784453
C 153	16.8	84.0	435 2	BB728884	BB728884	C 226	16.8	84.0	522 2	BB784453
C 154	16.8	84.0	435 4	BB728884	BB728884	C 227	16.8	84.0	525 4	BB158096
C 155	16.8	84.0	435 4	BB728884	BB728884	C 228	16.8	84.0	525 4	BB158096
C 156	16.8	84.0	435 4	BB728884	BB728884	C 229	16.8	84.0	525 5	BO748583
C 157	16.8	84.0	436 2	BB133728	BB133728	C 230	16.8	84.0	525 5	BO748583
C 158	16.8	84.0	436 2	BB133728	BB133728	C 231	16.8	84.0	525 7	CK340960
C 159	16.8	84.0	436 7	CK340892	CK340892	C 232	16.8	84.0	525 7	CK340960
C 160	16.8	84.0	436 7	CK340892	CK340892	C 233	16.8	84.0	526 8	AZ120597
C 161	16.8	84.0	438 6	CD802736	CD802736	C 234	16.8	84.0	526 8	AZ120597
C 162	16.8	84.0	438 6	CD802736	CD802736	C 235	16.8	84.0	527 8	AZ461633
C 163	16.8	84.0	440 8	AZ379582	AZ379582	C 236	16.8	84.0	527 8	AZ461633
C 164	16.8	84.0	440 8	AZ379582	AZ379582	C 237	16.8	84.0	528 2	BB752621
C 165	16.8	84.0	441 2	AA762330	AA762330	C 238	16.8	84.0	528 2	BB752621
C 166	16.8	84.0	441 2	AA762330	AA762330	C 239	16.8	84.0	533 5	BX524773
C 167	16.8	84.0	443 2	BB159594	BB159594	C 240	16.8	84.0	533 5	BX524773
C 168	16.8	84.0	443 2	BB159594	BB159594	C 241	16.8	84.0	533 8	AZ123002
C 169	16.8	84.0	456 1	AA267947	AA267947	C 242	16.8	84.0	533 8	AZ123002
C 170	16.8	84.0	456 1	AA267947	AA267947	C 243	16.8	84.0	536 8	BH031963

C 244	16.8	84.0	536	8	BH031963	317	16.8	84.0	626	9	CR082848
C 245	16.8	84.0	537	1	AA389154	C 318	16.8	84.0	626	9	CR082848 Reverse B
C 246	16.8	84.0	537	1	AA389154	C 319	16.8	84.0	629	8	AZ113897
C 247	16.8	84.0	537	9	CR264982	C 320	16.8	84.0	629	8	AZ113897
C 248	16.8	84.0	537	9	CR264982	C 321	16.8	84.0	630	7	CP906196
C 249	16.8	84.0	539	8	AZ264310	C 322	16.8	84.0	630	7	CP906196
C 250	16.8	84.0	539	8	AZ264310	C 323	16.8	84.0	630	7	CP906196
C 251	16.8	84.0	541	8	AZ651848	C 324	16.8	84.0	630	8	BH065057
C 252	16.8	84.0	541	8	AZ651848	C 325	16.8	84.0	630	8	BH065057
C 253	16.8	84.0	543	7	CK340918	C 326	16.8	84.0	632	2	BE281764
C 254	16.8	84.0	543	7	CK340918	C 327	16.8	84.0	632	8	AZ792596
C 255	16.8	84.0	543	7	CK341119	C 328	16.8	84.0	632	8	AZ792596
C 256	16.8	84.0	543	7	CK341119	C 329	16.8	84.0	634	8	AZ494009
C 257	16.8	84.0	545	8	AZ889925	C 330	16.8	84.0	634	8	AZ494009
C 258	16.8	84.0	548	9	CR141987	C 331	16.8	84.0	640	8	AZ800701
C 259	16.8	84.0	548	9	CR141987	C 332	16.8	84.0	640	8	AZ800701
C 260	16.8	84.0	555	2	BE913551	C 333	16.8	84.0	641	8	AZ505060
C 261	16.8	84.0	555	2	BE913551	C 334	16.8	84.0	641	8	AZ505060
C 262	16.8	84.0	555	4	BG147317	C 335	16.8	84.0	641	8	AZ999640
C 263	16.8	84.0	557	4	BG147317	C 336	16.8	84.0	641	8	AZ999640
C 264	16.8	84.0	557	4	BG147317	C 337	16.8	84.0	644	2	BF148642
C 265	16.8	84.0	559	1	AA110628	C 338	16.8	84.0	644	2	BF148642
C 266	16.8	84.0	559	1	AA110628	C 339	16.8	84.0	649	2	CR09325
C 267	16.8	84.0	559	8	AZ367994	C 340	16.8	84.0	649	9	CR09325
C 268	16.8	84.0	559	8	AZ367994	C 341	16.8	84.0	650	8	AQ985141
C 269	16.8	84.0	560	4	BG067777	C 342	16.8	84.0	650	8	AQ985141
C 270	16.8	84.0	560	4	BG067777	C 343	16.8	84.0	650	8	AZ241609
C 271	16.8	84.0	568	8	AZ336285	C 344	16.8	84.0	650	8	AZ241609
C 272	16.8	84.0	568	8	AZ336285	C 345	16.8	84.0	650	8	AZ410608
C 273	16.8	84.0	571	6	CA877937	C 346	16.8	84.0	650	8	AZ410608
C 274	16.8	84.0	571	6	CA877937	C 347	16.8	84.0	651	8	AZ609056
C 275	16.8	84.0	575	6	CD876173	C 348	16.8	84.0	651	8	AZ609056
C 276	16.8	84.0	575	6	CD876173	C 349	16.8	84.0	655	8	AZ764190
C 277	16.8	84.0	576	8	AZ790675	C 350	16.8	84.0	655	8	AZ764190
C 278	16.8	84.0	576	8	AZ790675	C 351	16.8	84.0	657	8	AZ791482
C 279	16.8	84.0	578	8	AZ587544	C 352	16.8	84.0	657	8	AZ791482
C 280	16.8	84.0	578	8	AZ587544	C 353	16.8	84.0	657	8	AZ950551
C 281	16.8	84.0	580	8	AZ473435	C 354	16.8	84.0	657	8	AZ950551
C 282	16.8	84.0	580	8	AZ473435	C 355	16.8	84.0	658	8	AZ652319
C 283	16.8	84.0	581	9	CR262174	C 356	16.8	84.0	658	8	AZ652319
C 284	16.8	84.0	581	9	CR262174	C 357	16.8	84.0	660	8	AZ619777
C 285	16.8	84.0	584	8	AZ327480	C 358	16.8	84.0	660	8	AZ619777
C 286	16.8	84.0	584	8	AZ327480	C 359	16.8	84.0	661	1	AV303415
C 287	16.8	84.0	587	8	AZ355676	C 360	16.8	84.0	661	1	AV303415
C 288	16.8	84.0	587	8	AZ355676	C 361	16.8	84.0	661	8	AZ450945
C 289	16.8	84.0	588	8	AQ934840	C 362	16.8	84.0	661	8	AZ450945
C 290	16.8	84.0	588	8	AQ934840	C 363	16.8	84.0	663	9	CR276027
C 291	16.8	84.0	590	6	CA879114	C 364	16.8	84.0	663	9	CR276027
C 292	16.8	84.0	590	6	CA879114	C 365	16.8	84.0	665	2	BE287465
C 293	16.8	84.0	590	8	AZ326978	C 366	16.8	84.0	665	2	BE287465
C 294	16.8	84.0	593	8	AZ326978	C 367	16.8	84.0	672	8	AZ581892
C 295	16.8	84.0	593	8	AZ326978	C 368	16.8	84.0	672	8	AZ581892
C 296	16.8	84.0	594	2	BF143384	C 369	16.8	84.0	684	8	AZ970909
C 297	16.8	84.0	594	2	BF143384	C 370	16.8	84.0	684	8	AZ970909
C 298	16.8	84.0	594	2	BF143384	C 371	16.8	84.0	687	6	BY748643
C 299	16.8	84.0	595	8	AZ903084	C 372	16.8	84.0	687	6	BY748643
C 300	16.8	84.0	595	8	AZ903084	C 373	16.8	84.0	691	9	CR134687
C 301	16.8	84.0	600	5	BU919589	C 374	16.8	84.0	691	9	CR134687
C 302	16.8	84.0	600	5	BU919589	C 375	16.8	84.0	693	8	AZ503640
C 303	16.8	84.0	600	8	AZ609157	C 376	16.8	84.0	693	8	AZ503640
C 304	16.8	84.0	600	8	AZ609157	C 377	16.8	84.0	693	9	BE962618
C 305	16.8	84.0	605	5	BQ268631	C 378	16.8	84.0	693	9	BE962618
C 306	16.8	84.0	605	5	BQ268631	C 379	16.8	84.0	694	8	AZ989689
C 307	16.8	84.0	609	8	AZ972769	C 380	16.8	84.0	694	8	AZ989689
C 308	16.8	84.0	609	8	AZ972769	C 381	16.8	84.0	697	8	AZ271298
C 309	16.8	84.0	619	8	AZ871987	C 382	16.8	84.0	697	8	AZ271298
C 310	16.8	84.0	619	8	AZ871987	C 383	16.8	84.0	699	8	AZ707813
C 311	16.8	84.0	622	8	AZ947859	C 384	16.8	84.0	699	8	AZ707813
C 312	16.8	84.0	622	8	AZ947859	C 385	16.8	84.0	699	8	AZ982053
C 313	16.8	84.0	625	5	BP765720	C 386	16.8	84.0	699	8	AZ982053
C 314	16.8	84.0	625	5	BP765720	C 387	16.8	84.0	707	4	BI105779
C 315	16.8	84.0	626	2	BF168283	C 388	16.8	84.0	707	4	BI105779
C 316	16.8	84.0	626	2	BF168283	C 389	16.8	84.0	708	8	AZ379355

C 390	16.8	84.0	708	8	A2379355
C 391	16.8	84.0	708	8	A2843347
C 392	16.8	84.0	708	8	A2843347
C 393	16.8	84.0	710	2	BP162719
C 394	16.8	84.0	710	2	BP162719
C 395	16.8	84.0	714	2	AG578809
C 396	16.8	84.0	714	2	AG578809
C 397	16.8	84.0	717	9	CR218886
C 398	16.8	84.0	717	9	CR218886
C 399	16.8	84.0	731	7	CNS25792
C 400	16.8	84.0	731	7	CNS25792
C 401	16.8	84.0	735	4	BI158715
C 402	16.8	84.0	735	4	BI158715
C 403	16.8	84.0	735	4	CR053295
C 404	16.8	84.0	735	4	CR053295
C 405	16.8	84.0	738	9	CNS0104M
C 406	16.8	84.0	738	9	CNS0104M
C 407	16.8	84.0	742	9	AG469535
C 408	16.8	84.0	742	9	AG469535
C 409	16.8	84.0	747	9	AG421780
C 410	16.8	84.0	747	9	AG421780
C 411	16.8	84.0	751	9	AG563398
C 412	16.8	84.0	751	9	AG563398
C 413	16.8	84.0	765	9	AG497680
C 414	16.8	84.0	765	9	AG497680
C 415	16.8	84.0	773	7	CO041908
C 416	16.8	84.0	773	7	CO041908
C 417	16.8	84.0	774	9	AG503294
C 418	16.8	84.0	774	9	AG503294
C 419	16.8	84.0	774	9	CR124419
C 420	16.8	84.0	774	9	CR124419
C 421	16.8	84.0	782	9	AG477284
C 422	16.8	84.0	782	9	AG477284
C 423	16.8	84.0	785	9	CR163285
C 424	16.8	84.0	785	9	CR163285
C 425	16.8	84.0	791	9	CR233592
C 426	16.8	84.0	791	9	CR233592
C 427	16.8	84.0	792	9	CR218103
C 428	16.8	84.0	792	9	CR218103
C 429	16.8	84.0	794	9	CR049484
C 430	16.8	84.0	794	9	CR049484
C 431	16.8	84.0	798	9	CR024249
C 432	16.8	84.0	798	9	CR024249
C 433	16.8	84.0	799	9	AG532346
C 434	16.8	84.0	799	9	AG532346
C 435	16.8	84.0	801	9	CR032079
C 436	16.8	84.0	801	9	CR032079
C 437	16.8	84.0	801	9	CR186213
C 438	16.8	84.0	806	9	CR271667
C 439	16.8	84.0	806	9	CR271667
C 440	16.8	84.0	815	9	CR209043
C 441	16.8	84.0	815	9	CR209043
C 442	16.8	84.0	815	9	CR209043
C 443	16.8	84.0	817	9	CR209043
C 444	16.8	84.0	817	9	CR209043
C 445	16.8	84.0	824	8	BH369337
C 446	16.8	84.0	824	8	BH369337
C 447	16.8	84.0	830	9	CR053794
C 448	16.8	84.0	830	9	CR053794
C 449	16.8	84.0	832	9	CR053794
C 450	16.8	84.0	832	9	CR053794
C 451	16.8	84.0	837	9	CR206994
C 452	16.8	84.0	837	9	CR206994
C 453	16.8	84.0	842	9	CR151241
C 454	16.8	84.0	842	9	CR151241
C 455	16.8	84.0	847	9	CR245141
C 456	16.8	84.0	847	9	CR245141
C 457	16.8	84.0	858	9	CR068569
C 458	16.8	84.0	858	9	CR068569
C 459	16.8	84.0	862	9	CR087761
C 460	16.8	84.0	862	9	CR087761
C 461	16.8	84.0	871	9	CR232774
C 462	16.8	84.0	871	9	CR232774

A2379355	IM0134F05
A2843347	2M0142L08
A2843347	2M0142L08
BP162719	601769374
BP162719	601769374
AG578809	Mus muscu
AG578809	Mus muscu
CR218886	Forward s
CR218886	Forward s
CNS25792	UI-M-HNO-
CNS25792	UI-M-HNO-
BI158715	602921796
BI158715	602921796
CR053295	Reverse s
CR053295	Reverse s
CR053295	Reverse s
AL155654	Anopheles
AL155654	Anopheles
AG469535	Mus muscu
AG469535	Mus muscu
AG421780	Mus muscu
AG421780	Mus muscu
AG563398	Mus muscu
AG563398	Mus muscu
AG497680	Mus muscu
AG497680	Mus muscu
CO041908	UI-M-FY0-
CO041908	UI-M-FY0-
AG503294	Mus muscu
AG503294	Mus muscu
CR124419	Forward s
CR124419	Forward s
AG477284	Mus muscu
AG477284	Mus muscu
CR163285	Reverse s
CR163285	Reverse s
CR233592	Reverse s
CR233592	Reverse s
CR218103	Forward s
CR218103	Forward s
CR049484	Reverse s
CR049484	Reverse s
CR024249	Forward s
CR024249	Forward s
AG532346	Mus muscu
AG532346	Mus muscu
CR032079	Forward s
CR032079	Forward s
CR186213	Reverse s
CR186213	Reverse s
CR271667	Reverse s
CR271667	Reverse s
CR209043	Reverse s
CR209043	Reverse s
CR209043	Reverse s
BH369337	AG-ND-173
BH369337	AG-ND-173
CR053794	Reverse s
CR053794	Reverse s
CR053794	Reverse s
BX997284	Reverse s
BX997284	Reverse s
CR206994	Forward s
CR206994	Forward s
CR151241	Forward s
CR151241	Forward s
CR245141	Reverse s
CR245141	Reverse s
CR068569	Forward s
CR068569	Forward s
CR087761	Forward s
CR087761	Forward s
CR232774	Forward s
CR232774	Forward s

C 463	16.8	84.0	873	4	BI152573
C 464	16.8	84.0	873	4	BI152573
C 465	16.8	84.0	874	5	BX330700
C 466	16.8	84.0	874	5	BX330700
C 467	16.8	84.0	891	6	CR193342
C 468	16.8	84.0	891	6	CR193342
C 469	16.8	84.0	894	9	CR253312
C 470	16.8	84.0	894	9	CR253312
C 471	16.8	84.0	900	9	CR201245
C 472	16.8	84.0	900	9	CR201245
C 473	16.8	84.0	908	9	CNS0104M
C 474	16.8	84.0	908	9	CNS0104M
C 475	16.8	84.0	911	2	BE367834
C 476	16.8	84.0	911	2	BE367834
C 477	16.8	84.0	918	9	CR216021
C 478	16.8	84.0	918	9	CR216021
C 479	16.8	84.0	923	9	CR181912
C 480	16.8	84.0	923	9	CR181912
C 481	16.8	84.0	924	9	CR063036
C 482	16.8	84.0	924	9	CR063036
C 483	16.8	84.0	930	9	BX966088
C 484	16.8	84.0	930	9	BX966088
C 485	16.8	84.0	961	9	CR027287
C 486	16.8	84.0	961	9	CR027287
C 487	16.8	84.0	1059	7	CF580880
C 488	16.8	84.0	1059	7	CF580880
C 489	16.8	84.0	1077	7	CK163293
C 490	16.8	84.0	1077	7	CK163293
C 491	16.8	84.0	1250	9	AG310691
C 492	16.8	84.0	1250	9	AG310691
C 493	16.8	84.0	1688	3	AK080131
C 494	16.8	84.0	1688	3	AK080131
C 495	16.8	84.0	2422	3	AK042011
C 496	16.8	84.0	2422	3	AK042011
C 497	16.8	84.0	2742	3	AK030879
C 498	16.8	84.0	2742	3	AK030879
C 499	16.4	82.0	192	2	BB416909
C 500	16.4	82.0	192	2	BB416909

## ALIGNMENTS

RESULT 1  
CR249440  
LOCUS  
DEFINITION Forward strand read from insert in 5'HRT insertion targeting and  
chromosome engineering clone MHPN63p13, genomic survey sequence.  
ACCESSION CR249440  
VERSION GI:50028293  
KEYWORDS GSS; genome survey sequence; MICR.  
SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
(bases 1 to 683)

## REFERENCE

Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,J.,  
Jenkins,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.  
Direct Submission  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
UK. http://www.sanger.ac.uk/MICR

## AUTHORS

ADAMS,D.J., BIGGS,P.J., COX,A.V., DAVIES,R.M., VAN DER WEYDEN,J.,  
JENKINS,J., SMITH,J., PLUMB,R.W., TAYLOR,R.G., NISHIJIMA,I., YU,Y.,  
ROGERS,J. and BRADLEY,A.  
Direct Submission  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
UK. http://www.sanger.ac.uk/MICR

## FEATURES

LOCATION/Qualifiers  
1..683  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPN63p13"  
/clone\_1b="MHPN"

## ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 683;  
Best Local Similarity 95.0%; Pred. No. 2.8e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 DB 172 GGACATGCCCGGCGCATGTCC 191

RESULT 2  
 CR249440/c 683 bp DNA linear GSS 06-JUL-2004  
 LOCUS CR249440/c  
 DEFINITION Forward strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN63p13, genomic survey sequence.

ACCESSION CR249440  
 VERSION CR249440.1 GI:50028293  
 KEYWORDS GSS; genome survey sequence; MICEP.  
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 683)  
 Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jomkero,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.  
 Direct Submission  
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICEP  
 Location/Qualifiers  
 1..683  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHPN63p13"  
 /clone\_lib="MHPN"

ORIGIN  
 Query Match 92.0%; Score 18.4; DB 9; Length 683;  
 Best Local Similarity 95.0%; Pred. No. 2.8e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 DB 191 GGACCTGCCCGGCGCATGTCC 172

RESULT 3  
 BH781927 409 bp DNA linear GSS 28-MAR-2002  
 LOCUS fzm0001f001g06k0 fzm filtered library Zea mays genomic clone  
 DEFINITION fzm0001f001g06 5', genomic survey sequence.  
 ACCESSION BH781927  
 VERSION BH781927.1 GI:19785173  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 409)  
 Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.  
 Genethresher methylation filtered genomic sequences from maize  
 Unpublished (2002)  
 Contact: Bedell JA  
 Orion Genomics, LLC  
 4041 Forest Park Ave, St. Louis, MO 63108, USA  
 Tel: 314 615 6979  
 Fax: 314 615 5975  
 Email: jbedell@oriongenomics.com  
 Plate: fzm0001f001 row: g column: 06  
 Seq primer: SK reverse  
 Classes: shotgun  
 High quality sequence stop: 409.  
 Location/Qualifiers  
 1..409

FEATURES  
 source

/organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="Mo17"  
 /db\_xref="taxon:4577"  
 /clone="fzm0001f001g06"  
 /clone\_lib="fzm filtered library"  
 /note="Organ: leaf; Vector: pBCSK(-); Site\_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells."

## ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 409;  
 Best Local Similarity 94.7%; Pred. No. 8.5e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGCATGTCC 20  
 |||||  
 DB 45 GACATGCCCGGCGCATGTCC 63

RESULT 4  
 BH781927/c 409 bp DNA linear GSS 28-MAR-2002  
 LOCUS fzm0001f001g06k0 fzm filtered library Zea mays genomic clone  
 DEFINITION fzm0001f001g06 5', genomic survey sequence.  
 ACCESSION BH781927  
 VERSION BH781927.1 GI:19785173  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 409)  
 Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.  
 Genethresher methylation filtered genomic sequences from maize  
 Unpublished (2002)  
 Contact: Bedell JA  
 Orion Genomics, LLC  
 4041 Forest Park Ave, St. Louis, MO 63108, USA  
 Tel: 314 615 6979  
 Fax: 314 615 5975  
 Email: jbedell@oriongenomics.com  
 Plate: fzm0001f001 row: g column: 06  
 Seq primer: SK reverse  
 Classes: shotgun  
 High quality sequence stop: 409.  
 Location/Qualifiers  
 1..409  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="Mo17"  
 /db\_xref="taxon:4577"  
 /clone="fzm0001f001g06"  
 /clone\_lib="fzm filtered library"  
 /note="Organ: leaf; Vector: pBCSK(-); Site\_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells."

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source

Query Match 87.0%; Score 17.4; DB 8; Length 409;  
 Best Local Similarity 94.7%; Pred. No. 8.5e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 19  
 |||||  
 DB 63 GGACATGCCCGGCGCATGTCC 45

FEATURES  
 source

RESULT 5  
 BG556896 482 bp mRNA linear EST 10-APR-2001  
 LOCUS EML\_40\_A07.bl\_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA  
 DEFINITION  
 ACCESSION BG556896  
 VERSION BG556896.1 GI:13585894  
 KEYWORDS EST.  
 SOURCE Sorghum bicolor (sorghum)  
 ORGANISM Sorghum bicolor  
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 482)  
 Reid,S.P., Cordonnier-Pratt,M.-M., Gingie,A. and Pratt,L.H.  
 An EST database from Sorghum: developing embryos  
 Unpublished (2000)  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
 Seq primer: JEN REV  
 High quality sequence stop: 480  
 POLYA=No.

FEATURES  
 source 1..482  
 location/Qualifiers  
 /organism="Sorghum bicolor"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Embryo 1 (EM1)"  
 /note="Organ: Embryos germinated for 24 hr; Vector: pBluescript II from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN  
 Query Match 87.0%; Score 17.4; DB 4; Length 482;  
 Best Local Similarity 94.7%; Pred. No. 8.5e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 GGCATGCCCGGGCATGTC 19  
 |||||  
 Db 247 GGACACGCCCGGGCATGTC 265

RESULT 6  
 BG556896 482 bp mRNA linear EST 10-APR-2001  
 LOCUS EML\_40\_A07.bl\_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA  
 DEFINITION  
 ACCESSION BG556896  
 VERSION BG556896.1 GI:13585894  
 KEYWORDS EST.  
 SOURCE Sorghum bicolor (sorghum)  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 482)  
 Reid,S.P., Cordonnier-Pratt,M.-M., Gingie,A. and Pratt,L.H.  
 An EST database from Sorghum: developing embryos  
 Unpublished (2000)  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
 Seq primer: JEN REV  
 High quality sequence stop: 480  
 POLYA=No.

FEATURES  
 source 1..482  
 location/Qualifiers  
 /organism="Sorghum bicolor"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Embryo 1 (EM1)"  
 /note="Organ: Embryos germinated for 24 hr; Vector: pBluescript II from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN  
 Query Match 87.0%; Score 17.4; DB 4; Length 482;  
 Best Local Similarity 94.7%; Pred. No. 8.5e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GACATGCCCGGGCATGCC 20  
 |||||  
 Db 265 GACATGCCCGGGCATGCC 247

RESULT 7  
 BI098283 557 bp mRNA linear EST 26-JUN-2001  
 LOCUS IPl\_30\_D12.bl\_A002 Immature panicle 1 (IPl) Sorghum bicolor cDNA, mRNA sequence.  
 DEFINITION  
 ACCESSION BI098283  
 VERSION BI098283.1 GI:14569865  
 KEYWORDS EST.  
 SOURCE Sorghum bicolor (sorghum)  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 557)  
 Klein,R.R., Cordonnier-Pratt,M.-M., Gingie,A., Sudman,M. and Pratt,L.H.  
 An EST database from Sorghum: developing preanthesis panicles  
 Unpublished (2001)  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyTmX or T7 sequencing primer, are presented as the reverse complement.  
 Seq primer: JEN REV  
 High quality sequence stop: 551  
 POLYA=No.

FEATURES  
 source 1..557  
 location/Qualifiers  
 /organism="Sorghum bicolor"  
 /mol\_type="mRNA"  
 /cultivar="Brx623"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Immature panicle 1 (IPl)"  
 /note="Organ: Developing preanthesis panicles; Vector: pBluescript II SK(-) from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the



cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

## ORIGIN

Query Match 87.0%; Score 17.4; DB 4; Length 557;  
Best Local Similarity 94.7%; Pred. No. 8.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGATGTC 19  
|||||  
DB 75 GGACATGCCCGGCGATGTC 93

RESULT 8  
BI098283/c  
LOCUS  
DEFINITION  
IPI\_30\_D12.b1\_A002 Immature panicle 1 (IPI) Sorghum bicolor cDNA,  
mRNA sequence.  
ACCESSION  
BI098283  
VERSION  
BI098283.1 GI:14569865  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Sorghum bicolor (sorghum)

REFERENCE  
AUTHORS  
Klein, R.R., Cordomier-Pratt, M.-M., Gingle, A., Sudman, M. and  
Pratt, L.H.  
An EST database from Sorghum: developing preanthesis panicles  
Unpublished (2001)  
Contact: Cordomier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

TITLE  
JOURNAL  
COMMENT  
Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for high quality sequence is  
20. Three-prime sequences, which are obtained with PolyMix or T7  
Sequencing primer, are presented as the reverse complement.  
Seq primer: JEN REV  
High quality sequence stop: 551  
POLYA=No.

FEATURES  
source  
1. .557  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="BTx623"  
/db\_xref="taxon:4558"  
/clone\_lib="Immature panicle 1 (IPI)"  
/note="Organ: Developing preanthesis panicles; Vector:  
pBluescript II SK(-) from lambda ZAP II; Site 1: XhoI;  
Site 2: EcoRI; The library was made from poly-A RNA in the  
cloning vector lambda ZAP II. Clones to be sequenced were  
prepared by mass excision."

## ORIGIN

Query Match 87.0%; Score 17.4; DB 4; Length 557;  
Best Local Similarity 94.7%; Pred. No. 8.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGATGTC 20  
|||||  
DB 93 GACATGCCCGGCGATGTC 75

RESULT 9  
CG270390  
LOCUS  
DEFINITION  
CG270390 ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0783J21,  
genomic survey sequence.  
757 bp DNA linear GSS 25-AUG-2003

ACCESSION  
CG270390  
VERSION  
CG270390.1 GI:34182531  
KEYWORDS  
GSS.  
SOURCE  
Zea mays

REFERENCE  
AUTHORS  
Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 757)

TITLE  
JOURNAL  
COMMENT  
Unpublished (2002)  
Other GSSs: CG3DQ59TH  
Contact: Cathy Whitelaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TP  
Class: sheared ends.

FEATURES  
source  
1. .757  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZM 0.7.1.5 KB"  
/note="Vector: pBSCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 757;  
Best Local Similarity 94.7%; Pred. No. 8.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGATGTC 20  
|||||  
DB 243 GACATGCCCGGCGATGTC 261

RESULT 10  
CG270390  
LOCUS  
DEFINITION  
CG270390 ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0783J21,  
genomic survey sequence.  
ACCESSION  
CG270390  
VERSION  
CG270390.1 GI:34182531  
KEYWORDS  
GSS.  
SOURCE  
Zea mays

REFERENCE  
AUTHORS  
Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 757)

TITLE  
JOURNAL  
COMMENT  
Unpublished (2002)  
Other GSSs: CG3DQ59TH  
Contact: Cathy Whitelaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TP  
Class: sheared ends.  
Location/Qualifiers

```

source
1. .757
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lib="ZM07151G21"
/clone_1lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match      87.0%; Score 17.4; DB 9; Length 757;
Best Local Similarity 94.7%; Pred. No. 8.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 19
Db 261 GGACACGCCCGGCGCATGTC 243

RESULT 11
CG290290      843 bp   DNA      linear   GSS 25-AUG-2003
LOCUS
DEFINITION   CG2BC47TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBMA0751G21,
ACCESSION    CG290290
VERSION      CG290290.1 GI:34204504
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 843)
Whiteclaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: CG2BC47TV
Contact: Cathy Whiteclaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteclaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
location/Qualifiers
1..843
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lib="ZMMBMA0751G21"
/clone_1lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match      87.0%; Score 17.4; DB 9; Length 843;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGCATGTC 20
Db 524 GACATGCCCGGCGCATGTC 542

RESULT 12
CG290290      843 bp   DNA      linear   GSS 25-AUG-2003
LOCUS
DEFINITION   CG2BC47TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBMA0751G21,

```

```

genomic survey sequence.
ACCESSION    CG290290
VERSION      CG290290.1 GI:34204504
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 843)
Whiteclaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: CG2BC47TV
Contact: Cathy Whiteclaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteclaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
location/Qualifiers
1..843
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lib="ZMMBMA0751G21"
/clone_1lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match      87.0%; Score 17.4; DB 9; Length 843;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 19
Db 542 GGACACGCCCGGCGCATGTC 524

RESULT 13
CG236275      870 bp   DNA      linear   GSS 22-AUG-2003
LOCUS
DEFINITION   OG1CU67TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBMA0733K13,
ACCESSION    CG236275
VERSION      CG236275.1 GI:34136161
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 870)
Whiteclaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG1CU67TV
Contact: Cathy Whiteclaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteclaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
location/Qualifiers
1..870
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lib="ZMMBMA0733K13"
/clone_1lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match      87.0%; Score 17.4; DB 9; Length 843;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 19
Db 542 GGACACGCCCGGCGCATGTC 524

```

FEATURES  
SOURCE

Location/Qualifiers  
1..870  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_1lb="ZM 0.7-1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 870;  
Best Local Similarity 94.7%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTC 19  
DB 774 GGACACGCCCGGGCATGTC 792

RESULT 14  
CG236275/c

LOCUS CG236275 870 bp DNA linear GSS 22-AUG-2003

DEFINITION OG1CU677H ZM 0.7-1.5\_KB Zea mays genomic clone ZMMEMa0733K13,  
genomic survey sequence.

ACCESSION CG236275

VERSION CG236275.1 GI:34136161

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 870)  
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Cleck,R.W., Numborg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)

TITLE JOURNAL

COMMENT Other\_GSSs: OG1CU677V  
Contact: Cathy Whiteaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteaw@cigr.org  
Seq primer: 7R  
Class: sheared ends.

FEATURES  
SOURCE

Location/Qualifiers  
1..870  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_1lb="ZM 0.7-1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 870;  
Best Local Similarity 94.7%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGGCATGTC 20  
DB 792 GACATGCCCGGGCATGTC 774

RESULT 15  
B2964970

LOCUS B2964970 876 bp DNA linear GSS 25-MAR-2003

DEFINITION PUDGS23TD ZM 0.6-1.0\_KB Zea mays genomic clone ZMMBTa183D21,  
genomic survey sequence.

ACCESSION B2964970

VERSION B2964970.1 GI:29180280

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 876)  
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennerzen,J.  
Maize Genomics Consortium  
Unpublished (2003)

TITLE JOURNAL

COMMENT Contact: Cathy Whiteaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteaw@cigr.org  
Seq primer: 7P  
Class: sheared ends.

FEATURES  
SOURCE

Location/Qualifiers  
1..876  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_1lb="ZM 0.6-1.0\_KB"  
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
cot selected genomic DNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 876;  
Best Local Similarity 94.7%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTC 19  
DB 789 GGACACGCCCGGGCATGTC 807

RESULT 16  
B2964970/c

LOCUS B2964970 876 bp DNA linear GSS 25-MAR-2003

DEFINITION PUDGS23TD ZM 0.6-1.0\_KB Zea mays genomic clone ZMMBTa183D21,  
genomic survey sequence.

ACCESSION B2964970

VERSION B2964970.1 GI:29180280

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 876)  
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennerzen,J.  
Maize Genomics Consortium  
Unpublished (2003)

TITLE JOURNAL

COMMENT Contact: Cathy Whiteaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteaw@cigr.org  
Seq primer: 7P  
Class: sheared ends.

FEATURES  
SOURCE

Location/Qualifiers

```

source
1. 876
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM06183D21"
/clone_lib="ZM_0.6.1.0_KB"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN

Query Match      87.0%; Score 17.4; DB 8; Length 876;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GACATGCCCGGCATGTC 20
Db 807 GACATGCCCGGCATGTC 789

RESULT 17
LOCUS CL637471 877 bp DNA linear GSS 28-JUN-2004
DEFINITION CH243-SH5.SP6 CHORI-243 Ovis aries genomic clone CH243-SH5, genomic
survey sequence.
ACCESSION CL637471
VERSION CL637471.1 GI:49370767
KEYWORDS GSS.
SOURCE Ovis aries (sheep)
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Caprinae; Ovis.
1 (bases 1 to 877)
Zhaou S., Shetty J., de Jong P., McEwan J.C. and Oddy H.
Ovine BAC End Sequences from library CHORI-243
Unpublished (2004)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@tigr.org
Seq primer: SP6
Class: BAC ends.

FEATURES
source Location/Qualifiers
1..877
/organism="Ovis aries"
/mol_type="genomic DNA"
/strain="Texel breed"
/db_xref="taxon:9940"
/clone="CH243-SH5"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-243"
/notes="Vector: pPARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
The CHORI-243 sheep (M) (Ovis aries) BAC library produced
by Pieter de Jong's lab at CHORI
http://bacpac.chori.org/library.php?id=162"

ORIGIN

Query Match      87.0%; Score 17.4; DB 9; Length 877;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GACATGCCCGGCATGTC 19
Db 808 GGCATGCCCGGCATGTC 826

RESULT 18
LOCUS CL637471/c
```

```

LOCUS CL637471 877 bp DNA linear GSS 28-JUN-2004
DEFINITION CH243-SH5.SP6 CHORI-243 Ovis aries genomic clone CH243-SH5, genomic
survey sequence.
ACCESSION CL637471
VERSION CL637471.1 GI:49370767
KEYWORDS GSS.
SOURCE Ovis aries (sheep)
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Caprinae; Ovis.
1 (bases 1 to 877)
Zhaou S., Shetty J., de Jong P., McEwan J.C. and Oddy H.
Ovine BAC End Sequences from library CHORI-243
Unpublished (2004)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@tigr.org
Seq primer: SP6
Class: BAC ends.

FEATURES
source Location/Qualifiers
1..877
/organism="Ovis aries"
/mol_type="genomic DNA"
/strain="Texel breed"
/db_xref="taxon:9940"
/clone="CH243-SH5"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-243"
/notes="Vector: pPARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
The CHORI-243 sheep (M) (Ovis aries) BAC library produced
by Pieter de Jong's lab at CHORI
http://bacpac.chori.org/library.php?id=162"

ORIGIN

Query Match      87.0%; Score 17.4; DB 9; Length 877;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GACATGCCCGGCATGTC 20
Db 826 GACATGCCCGGCATGCC 808

RESULT 19
LOCUS CG236287 930 bp DNA linear GSS 22-AUG-2003
DEFINITION OG1CU677V_ZM_0.7.1.5_KB Zea mays genomic clone ZM06183K13,
genomic survey sequence.
ACCESSION CG236287
VERSION CG236287.1 GI:34136173
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 930)
Whitelaw C.A., Quackenbush J., Van Aken S., Uterback T.,
Reinick A., Fraser C.M., Budiman M.A., Bedell J.A., Rohlfing T.,
Citek R.W., Numborg A., Robbins D. and Lakey N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG1CU677H
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
```

Fax: 301-838-0208  
Email: white@tigr.org  
Seq primer: TP  
Classes: sheared ends.  
Location/Qualifiers  
1..930

FEATURES  
source  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZMMBMA0733K13"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 87.0%; Score 17.4; DB 9; Length 930;  
Best Local Similarity 94.7%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCGCGGCATGTC 20  
|||||  
Db 355 GACATGCGCGGCATGTC 373

RESULT 20 CG236287 930 bp DNA linear GSS 22-AUG-2003  
CG236287/c LOCUS  
DEFINITION OG2BC47TV ZM 0.7.1.5\_KB Zea mays genomic clone ZMMBMA0733K13,  
genomic survey sequence.  
ACCESSION CG236287  
VERSION CG236287.1 GI:34136173  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 930)

White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OG2BC47TV  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: white@tigr.org  
Seq primer: TP  
Class: sheared ends.

FEATURES  
source  
Location/Qualifiers  
1..930  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZMMBMA0733K13"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 87.0%; Score 17.4; DB 9; Length 930;  
Best Local Similarity 94.7%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATGCGCGGCATGTC 19  
|||||  
Db 373 GACATGCGCGGCATGTC 355

RESULT 21 CG290302 977 bp DNA linear GSS 25-AUG-2003  
CG290302 LOCUS  
DEFINITION OG2BC47TV ZM 0.7.1.5\_KB Zea mays genomic clone ZMMBMA0751G21,  
genomic survey sequence.

ACCESSION CG290302  
VERSION CG290302.1 GI:34204516  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 977)

White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OG2BC47TV  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: white@tigr.org  
Seq primer: TP  
Class: sheared ends.

FEATURES  
source  
Location/Qualifiers  
1..977  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZMMBMA0751G21"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 87.0%; Score 17.4; DB 9; Length 977;  
Best Local Similarity 94.7%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATGCGCGGCATGTC 19  
|||||  
Db 813 GACATGCGCGGCATGTC 831

RESULT 22 CG290302 977 bp DNA linear GSS 25-AUG-2003  
CG290302/c LOCUS  
DEFINITION OG2BC47TV ZM 0.7.1.5\_KB Zea mays genomic clone ZMMBMA0751G21,  
genomic survey sequence.

ACCESSION CG290302  
VERSION CG290302.1 GI:34204516  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 977)

White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OG2BC47TV  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@igr.org  
 Seq primer: TF  
 Class: sheared ends.  
 Location/Qualifiers

FEATURES  
 source  
 1..977  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMMA0751G21"  
 /clone\_1ib="ZM\_0.7-1.5\_KB"  
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

ORIGIN  
 Query Match 87.0%; Score 17.4; DB 9; Length 977;  
 Best Local Similarity 94.7%; Pred. No. 8.8e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GACATGCCCGGCGCATGTC 20  
 Db 831 GACATGCCCGGCGCATGTC 813

RESULT 23  
 CL640722 179 bp mRNA linear GSS 30-JUN-2004  
 LOCUS M003D05 GATC Gene Trap Library GV05C03 Mus musculus cDNA clone  
 DEFINITION M003D05, mRNA sequence.  
 ACCESSION CL640722  
 VERSION CL640722.1 GI:49489169  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS Hansen, J., Floss, T., van Sloun, P., Fuchtbauer, E.M., Vauti, F.,  
 1 (bases 1 to 179)  
 Arnold, H.H., Schutgen, F., Wurst, W., Von Melchner, H. and Ruiz, P.  
 A large-scale, gene-driven mutagenesis approach for the functional  
 analysis of the mouse genome  
 Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)  
 JOURNAL MEDLINE  
 PUBMED 12904583  
 COMMENT Contact: GGTC  
 German Genetrap Consortium (GGTC)  
 Email: info@genetrap.de  
 Ubetageo gene trap. Sequence tag generated by 5'RACE. Additional  
 sequence information can be found at:  
 'http://genetrap.gsf.de/project/web\_new/database/result\_clone.html'  
 clone\_id=M003D05' ES cell line harboring insertion mutation of  
 target gene is available at:  
 'http://genetrap.gsf.de/project/web\_new/order\_clones/howtoorder.htm  
 1' Inhouse Sequence Identifier: 01806  
 Class: Gene Trap.

FEATURES  
 source  
 1..179  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 SV"  
 /db\_xref="taxon:10090"  
 /clone="M003D05"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /cell\_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"  
 /clone\_1ib="GGTC Gene Trap Library GV05C03"  
 /note="Vector: Ubetageo"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 9; Length 179;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 GGACATGCCCGGCGCATGTC 20  
 Db 96 GGCGTTGCCCGGCGCATGTC 115

RESULT 24  
 CL640722 179 bp mRNA linear GSS 30-JUN-2004  
 LOCUS M003D05 GATC Gene Trap Library GV05C03 Mus musculus cDNA clone  
 DEFINITION M003D05, mRNA sequence.  
 ACCESSION CL640722  
 VERSION CL640722.1 GI:49489169  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS Hansen, J., Floss, T., van Sloun, P., Fuchtbauer, E.M., Vauti, F.,  
 1 (bases 1 to 179)  
 Arnold, H.H., Schutgen, F., Wurst, W., Von Melchner, H. and Ruiz, P.  
 A large-scale, gene-driven mutagenesis approach for the functional  
 analysis of the mouse genome  
 Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)  
 JOURNAL MEDLINE  
 PUBMED 12904583  
 COMMENT Contact: GGTC  
 German Genetrap Consortium (GGTC)  
 Email: info@genetrap.de  
 Ubetageo gene trap. Sequence tag generated by 5'RACE. Additional  
 sequence information can be found at:  
 'http://genetrap.gsf.de/project/web\_new/database/result\_clone.html'  
 clone\_id=M003D05' ES cell line harboring insertion mutation of  
 target gene is available at:  
 'http://genetrap.gsf.de/project/web\_new/order\_clones/howtoorder.htm  
 1' Inhouse Sequence Identifier: 01806  
 Class: Gene Trap.

FEATURES  
 source  
 1..179  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 SV"  
 /db\_xref="taxon:10090"  
 /clone="M003D05"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /cell\_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"  
 /clone\_1ib="GGTC Gene Trap Library GV05C03"  
 /note="Vector: Ubetageo"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 9; Length 179;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGACATGCCCGGCGCATGTC 20  
 Db 115 GGACATGCCCGGCGCATGTC 96

RESULT 25  
 B2691000 221 bp DNA linear GSS 14-FEB-2003  
 LOCUS M003D05 GV05C03 Mus musculus genomic clone M003D05, genomic survey  
 DEFINITION B2691000  
 ACCESSION B2691000  
 VERSION B2691000.1 GI:28382804  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 221)  
 AUTHORS Hansen, J., Floss, T., van Sloun, P., Fuchtbauer, B.M., Vauti, F., Arnold, H.H., Schuitgen, F., Wurst, W., Von Melchner, H. and Ruiz, P.  
 TITLE A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)  
 MEDLINE 22810117  
 PUBMED 12904583  
 COMMENT Contact: Wurst W  
 Institute of Developmental Genetics  
 GSF - National Research Center for Environment and Health  
 Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany  
 Tel: 49-89-3187-4110  
 Fax: 49-89-3187-3099  
 Email: wurst@gsf.de  
 U3btageo gene trap. Tag generated by RACE. Additional sequence information can be found at:  
 'http://genetrap.gsf.de/project/web\_new/database/result\_clone.html?clone\_id=M03D05'. ES cell line harboring insertion mutation of target gene is available at:  
 'http://genetrap.gsf.de/project/web\_new/order\_clones/howtoorder.htm'  
 1,  
 Inhouse Sequence Identifier: 01806  
 Class: Gene Trap.  
 FEATURES  
 source Location/Qualifiers  
 1..221  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129 Sv"  
 /db\_xref="taxon:10090"  
 /clone="M003D05"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /cell\_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"  
 /clone\_1lb="GV05C03"  
 /note="Vector: U3btageo"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 8; Length 221;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 65 GGACATGCCCGGCGCAAGCC 84

DB  
 BZ691000 221 bp DNA linear GSS 14-FEB-2003  
 BZ691000/C M003D05 GV05C03 Mus musculus genomic clone M003D05, genomic survey sequence.  
 ACCESSION BZ691000  
 VERSION BZ691000.1 GI:28382804  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 221)  
 Hansen, J., Floss, T., van Sloun, P., Fuchtbauer, B.M., Vauti, F., Arnold, H.H., Schuitgen, F., Wurst, W., Von Melchner, H. and Ruiz, P.  
 TITLE A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)  
 MEDLINE 22810117  
 PUBMED 12904583  
 COMMENT Contact: Wurst W  
 Institute of Developmental Genetics  
 GSF - National Research Center for Environment and Health  
 Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany  
 Tel: 49-89-3187-4110  
 Fax: 49-89-3187-3099

EMAIL: wurst@gsf.de  
 U3btageo gene trap. Tag generated by RACE. Additional sequence information can be found at:  
 'http://genetrap.gsf.de/project/web\_new/database/result\_clone.html?clone\_id=M03D05'. ES cell line harboring insertion mutation of target gene is available at:  
 'http://genetrap.gsf.de/project/web\_new/order\_clones/howtoorder.htm'  
 1,  
 Inhouse Sequence Identifier: 01806  
 Class: Gene Trap.  
 FEATURES  
 source Location/Qualifiers  
 1..221  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129 Sv"  
 /db\_xref="taxon:10090"  
 /clone="M003D05"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /cell\_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"  
 /clone\_1lb="GV05C03"  
 /note="Vector: U3btageo"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 8; Length 221;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCGCATGTCC 20  
 |||||  
 84 GGCGTTCGCCCGGCGCATGTCC 65

DB  
 AV322349 227 bp mRNA linear EST 09-NOV-1999  
 AV322349 RIKEN full-length enriched, 14 days embryo thymus Mus musculus cDNA clone 6130401D12 3' similar to D86726 Mouse mRNA for mMS5, mRNA sequence.  
 ACCESSION AV322349  
 VERSION AV322349.1 GI:6292237  
 KEYWORDS EST.  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 227)  
 Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kat, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomioka, N., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Kono, H., et al. 1999)  
 Unpublished (1999)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Shuhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsf.riken.jp, URL: http://genome.gsc.riken.jp/  
 Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsunaga, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 TITLE Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)  
 JOURNAL  
 COMMENT Itoh, M., Kiteunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

## FEATURES

source

## location/Qualifiers

1..227  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="6130401D12"  
/sex="mixed"  
/tissue\_type="thymus"  
/dev\_stage="14 days embryo"  
/lab\_host="DH10B"  
/clone\_1ib="RIKEN full-length enriched, 14 days embryo thymus"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCTCGAGTTCATTAATTAATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid pUC19(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 1; Length 227;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 GGACATGCCCGGCGCATGTC 20  
35 GGACAGCGCCGGCGCATTTCC 54

AV322349 227 bp mRNA linear EST 09-NOV-1999  
AV322349 RIKEN full-length enriched, 14 days embryo thymus Mus  
musculus cDNA clone 6130401D12 3' similar to D86726 Mouse mRNA for  
mMus5, mRNA sequence.

AV322349  
AV322349.1 GI:6292237

EST.  
Mus musculus (house mouse)

## KEYWORDS

Mus musculus

Mus musculus (house mouse)

## REFERENCE

AUTHORS

AV322349 227 bp mRNA linear EST 09-NOV-1999  
AV322349 RIKEN full-length enriched, 14 days embryo thymus Mus  
musculus cDNA clone 6130401D12 3' similar to D86726 Mouse mRNA for  
mMus5, mRNA sequence.  
AV322349  
AV322349.1 GI:6292237  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 227)  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,  
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirazane, T., Hori, F.,  
Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,  
Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,  
Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugihara, Y.,  
Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomioka, N.,  
Tsunoda, Y., Wataniki, A., Watanabe, S., Yamamura, T., Yasunishi, A.,

## TITLE

Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al. 1999)  
Unpublished (1999)

## JOURNAL

Contact: Yoshinide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Shukuro-cho, Tsukuba, Ibaraki, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: <http://genome.gsc.riken.jp/>

Sasaki, N., Izawa, M., Wataniki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,

Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y., and

Hayashizaki, Y.

Transcriptional sequencing: A method for DNA sequencing using RNA

polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

Itoh, M., Katsunari, T., Akiyama, U., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,

Okazaki, Y., and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

## FEATURES

source

## location/Qualifiers

1..227  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="6130401D12"  
/sex="mixed"  
/tissue\_type="thymus"  
/dev\_stage="14 days embryo"  
/lab\_host="DH10B"  
/clone\_1ib="RIKEN full-length enriched, 14 days embryo thymus"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCTCGAGTTCATTAATTAATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid pUC19(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 1; Length 227;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 GGACATGCCCGGCGCATGTC 20  
54 GGAATTCGCCGGCGCATTTCC 35

AV322349 227 bp DNA linear GSS 02-OCT-2000  
AV322349 RIKEN full-length enriched, 14 days embryo thymus Mus  
musculus cDNA clone 6130401D12 3' similar to D86726 Mouse mRNA for  
mMus5, mRNA sequence.

AV322349 227 bp DNA linear GSS 02-OCT-2000  
AV322349 RIKEN full-length enriched, 14 days embryo thymus Mus  
musculus cDNA clone 6130401D12 3' similar to D86726 Mouse mRNA for  
mMus5, mRNA sequence.



VERSION AZ377623.1 GI:10491323  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 269)  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,B., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contract: Robert B. Weiss  
 University of Utah Genome Center  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunne@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0132 row: L column: 01  
 Seq primer: CGTGTAAACGACGCGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 269.  
 Location/Qualifiers  
 1..269  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUC1M0132L01"  
 /sex="Male"  
 /lab\_host="B. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_1lb="Mouse 10kb plasmid UUC1M library"  
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent B. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 269;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 GGACATGCCCGGCGCATGTCC 20  
 |||  
 164 GGGCTGCGCCGGCGCATGTCC 183

RESULT 30  
 AZ377623/c 269 bp DNA linear GSS 02-OCT-2000  
 LOCUS 1M0132L01P Mouse 10kb plasmid UUC1M library Mus musculus genomic  
 DEFINITION clone UUC1M0132L01 F, genomic survey sequence.  
 ACCESSION AZ377623

VERSION AZ377623.1 GI:10491323  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 269)  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,B., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contract: Robert B. Weiss  
 University of Utah Genome Center  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunne@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0132 row: L column: 01  
 Seq primer: CGTGTAAACGACGCGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 269.  
 Location/Qualifiers  
 1..269  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUC1M0132L01"  
 /sex="Male"  
 /lab\_host="B. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_1lb="Mouse 10kb plasmid UUC1M library"  
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent B. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 269;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 GGACATGCCCGGCGCATGTCC 20  
 |||  
 183 GGACATGCCCGGCGCATGTCC 164

RESULT 31  
 AZ505789 274 bp DNA linear GSS 05-OCT-2000  
 LOCUS 1M0346E12R Mouse 10kb plasmid UUC1M library Mus musculus genomic  
 DEFINITION clone UUC1M0346E12 R, genomic survey sequence.  
 ACCESSION AZ505789

VERSION AZ505789.1 GI:10687105  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 274)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0346 row: E column: 12  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 274.  
location/Qualifiers  
1..274  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCGCM0346E12"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCGCM library"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 274;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCGCATGTCC 20  
DB 99 GGCGCTGCCCGGCGCATGTCC 118

RESULT 32  
AZ505789/c 274 bp DNA linear GSS 05-OCT-2000  
LOCUS IM0346E12R Mouse 10kb plasmid UUCGCM library Mus musculus genomic  
DEFINITION clone UUCGCM0346E12 R, genomic survey sequence.  
ACCESSION AZ505789

VERSION AZ505789.1 GI:10687105  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 274)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0346 row: E column: 12  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 274.  
location/Qualifiers  
1..274  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCGCM0346E12"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCGCM library"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 274;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCGCATGTCC 20  
DB 118 GGACATGCCCGGCGCATGTCC 99

RESULT 33  
BP090312 280 bp mRNA linear EST 30-JUN-2004  
LOCUS BP090312 Chlamydomonas reinhardtii C9 various conditions  
DEFINITION Chlamydomonas reinhardtii cDNA clone MX22h10\_r 5', mRNA sequence.  
ACCESSION BP090312

VERSION BP090312.1 GI:49462399  
 KEYWORDS EST.  
 SOURCE Chlamydomonas reinhardtii  
 ORGANISM Chlamydomonas reinhardtii  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.  
 REFERENCE 1 (bases 1 to 280)  
 AUTHORS Asamizu,E., Nakamura,Y., Mura,K., Fukuzawa,H., Fujiwara,S.,  
 Hitono,M., Iwamoto,K., Matsuda,Y., Minagawa,J., Shimogawara,K.,  
 Takahashi,Y. and Tabata,S.  
 TITLE Establishment of Publicly Available cDNA Material and Information  
 Resource of Chlamydomonas reinhardtii (Chlorophyta), to Facilitate  
 Gene Function Analysis  
 JOURNAL Phycologia (2004) In press  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..280  
 /organism="Chlamydomonas reinhardtii"  
 /mol\_type="mRNA"  
 /strain="C9"  
 /db\_xref="taxon:3055"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 5; Length 280;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 187 GGATGTGCGCGGCATGTC 206  
 1 GGACATGCCCGGCATGTC 20  
 |||||  
 187 GGATGTGCGCGGCATGTC 206

RESULT 34  
 BP090312/c  
 LOCUS BP090312 Chlamydomonas reinhardtii C9 various conditions  
 DEFINITION Chlamydomonas reinhardtii cDNA clone WX221h10\_r 5', mRNA sequence.  
 ACCESSION BP090312  
 VERSION BP090312.1 GI:49462399  
 KEYWORDS EST.  
 SOURCE Chlamydomonas reinhardtii  
 ORGANISM Chlamydomonas reinhardtii  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.  
 REFERENCE 1 (bases 1 to 280)  
 AUTHORS Asamizu,E., Nakamura,Y., Mura,K., Fukuzawa,H., Fujiwara,S.,  
 Hitono,M., Iwamoto,K., Matsuda,Y., Minagawa,J., Shimogawara,K.,  
 Takahashi,Y. and Tabata,S.  
 TITLE Establishment of Publicly Available cDNA Material and Information  
 Resource of Chlamydomonas reinhardtii (Chlorophyta), to Facilitate  
 Gene Function Analysis  
 JOURNAL Phycologia (2004) In press  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..280  
 /organism="Chlamydomonas reinhardtii"  
 /mol\_type="mRNA"  
 /strain="C9"  
 /db\_xref="taxon:3055"

/clone="WX221h10\_r"  
 /clone\_id="Chlamydomonas reinhardtii C9 various  
 conditions"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was made from a mixture of cells  
 grown under various conditions"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 5; Length 280;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 206 GGACATGCCCGGCATGTC 187  
 1 GGACATGCCCGGCATGTC 20  
 |||||  
 206 GGACATGCCCGGCATGTC 187

RESULT 35  
 BB093079  
 LOCUS BB093079 282 bp mRNA linear EST 23-JUN-2000  
 DEFINITION BB093079 RIKEN full-length enriched, 12 days embryo, embryonic body  
 between diaphragm region and neck Mus musculus cDNA clone  
 9430041B07 3', mRNA sequence.  
 ACCESSION BB093079  
 VERSION BB093079.1 GI:8676326  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 282)  
 AUTHORS Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,  
 Kiyosawa,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,  
 Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,  
 Shibata,K., Shibata,Y., Shigenoto,Y., Shingawa,A., Shiraki,T.,  
 Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,  
 Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watanishi,A.,  
 Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,  
 Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and  
 Hayashizaki,Y.  
 RIKEN Mouse ESTs (Kono,H., et al.)  
 Unpublished (2000)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Sasaki,N., Okazaki,Y., Westover,A., Itoh,M., Nagaoaka,S.,  
 Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Thermostabilization and thermocycling of thermostable enzymes by  
 chitosan and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh,M., Katsunari,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,  
 Okazaki,Y. and Hayashizaki,Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (http://genome.rtc.riken.go.jp) for  
 further details.  
 Location/Qualifiers  
 1..282  
 /organism="Mus musculus"  
 /mol\_type="mRNA"

FEATURES  
 source  
 Location/Qualifiers  
 1..282  
 /organism="Mus musculus"  
 /mol\_type="mRNA"

/db\_xref="taxon:10090"  
/clone="9430041E07"  
/tissue\_type="embryonic body between diaphragm region and neck"  
/dev\_stage="12 days embryo"  
/lab\_host="DH10B"  
/clone\_id="RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck"  
/note="Site 1: Sali; Site 2: BamHI, cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',  
GAGAGAGAGATCCCAAGACCTCTTTTCTTTTCTTTVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI."

ORIGIN  
Query Match 84.0%; Score 16.8; DB 2; Length 282;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
|||||  
36 GGACATGCCCGGCGCAAGCCC 55

RESULT 36  
LOCUS BB093079/c 282 bp mRNA linear EST 23-JUN-2000  
DEFINITION BB093079 RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck Mus musculus cDNA clone 9430041E07 3', mRNA sequence.  
ACCESSION BB093079  
VERSION BB093079.1 GI:8676326  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 282)  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Kono, H., et al.)  
JOURNAL Unpublished (2000)  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-ree@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoke, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

The most stabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Akiyama, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.  
location/Qualifiers

FEATURES  
source  
1..282  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="9430041E07"  
/tissue\_type="embryonic body between diaphragm region and neck"  
/dev\_stage="12 days embryo"  
/lab\_host="DH10B"  
/clone\_id="RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck"  
/note="Site 1: Sali; Site 2: BamHI, cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',  
GAGAGAGAGATCCCAAGACCTCTTTTCTTTTCTTTVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI."

ORIGIN  
Query Match 84.0%; Score 16.8; DB 2; Length 282;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
|||||  
55 GGCGTTGCCCGGCGCATGTCC 36

RESULT 37  
LOCUS BB157894 283 bp mRNA linear EST 29-JUN-2000  
DEFINITION BB157894 RIKEN full-length enriched, 16 days neonate thymus Mus musculus cDNA clone A130038G02 3', mRNA sequence.  
ACCESSION BB157894  
VERSION BB157894.1 GI:8613824  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 283)  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,



ORIGIN FLOC I."

Query Match 84.0%; Score 16.8; DB 2; Length 283;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||  
 57 GGCGCTTGCCCGGCGCATGTCC 38

RESULT 39  
 BB755436 283 bp mRNA linear EST 17-OCT-2001  
 BB755436 RIKEN full-length enriched, melanocyte Mus musculus cDNA  
 clone G270059B17 3', mRNA sequence.

DEFINITION  
 BB755436  
 BB755436.1 GI:16201995

ACCESSION  
 BB755436.1 GI:16201995

VERSION  
 EST.

KEYWORDS  
 Mus musculus (house mouse)

SOURCE  
 Mus musculus

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 283)

REFERENCE  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
 Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K.,  
 Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,  
 Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
 Saeki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,  
 Tanaka, T., Tomaru, A., Toya, T., Watanishi, A., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

TITLE  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

JOURNAL  
 Unpublished (2001)

COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

FEATURES  
 source  
 1. 283  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="G270059B17"  
 /cell\_type="melanocyte"

ORIGIN /clone\_lib="RIKEN full-length enriched, melanocyte"

Query Match 84.0%; Score 16.8; DB 2; Length 283;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
 |||  
 39 GGACATGCCCGGCGCATGTCC 58

RESULT 40  
 BB755436/c 283 bp mRNA linear EST 17-OCT-2001  
 BB755436 RIKEN full-length enriched, melanocyte Mus musculus cDNA  
 clone G270059B17 3', mRNA sequence.

DEFINITION  
 BB755436/c  
 BB755436/c.1 GI:16201995

ACCESSION  
 BB755436/c.1 GI:16201995

VERSION  
 EST.

KEYWORDS  
 Mus musculus (house mouse)

SOURCE  
 Mus musculus

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 283)

REFERENCE  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
 Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K.,  
 Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,  
 Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
 Saeki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,  
 Tanaka, T., Tomaru, A., Toya, T., Watanishi, A., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

TITLE  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

JOURNAL  
 Unpublished (2001)

COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

FEATURES  
 source  
 1. 283  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="G270059B17"  
 /cell\_type="melanocyte"

ORIGIN /clone\_lib="RIKEN full-length enriched, melanocyte"

Query Match 84.0%; Score 16.8; DB 2; Length 283;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
58 GGCGTTGGCCCGGCGCATGTCC 39

RESULT 41  
BY469224

DEFINITION BY469224 RIKEN full-length enriched, melanocyte Mus musculus cDNA  
clone G270064G22 3', mRNA sequence.

ACCESSION BY469224  
VERSION BY469224.1 GI:26803603

KEYWORDS EST.  
SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (bases 1 to 287)  
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Haegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsumura, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
Chochia, C., Corbani, L.B., Cousins, S., Dalla, E., Dragani, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gatiboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Matsuda, L., Marchionni, L., McKenzie, L., Mik, H., Nagashima, T.,  
Nunata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yangisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavanian, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, D., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
MEDLINE 22354683  
PUBMED 12466851  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Saitama-shi, Tsukuba-shi, Ibaraki, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
Alkawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,  
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numata, K.,  
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Watanabe, M. and  
Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

FEATURES  
source

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 287;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20  
44 GGCGTTGGCCCGGCGCATGTCC 63

RESULT 42  
BY469224/C  
LOCUS BY469224 RIKEN full-length enriched, melanocyte Mus musculus cDNA  
clone G270064G22 3', mRNA sequence.  
DEFINITION BY469224  
ACCESSION BY469224  
VERSION BY469224.1 GI:26803603  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 287)  
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Haegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsumura, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
Chochia, C., Corbani, L.B., Cousins, S., Dalla, E., Dragani, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Matsuda, L., Marchionni, L., McKenzie, L., Mik, H., Nagashima, T.,  
Nunata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multichannel sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Cells were provided by Drs. William J Pavan, Stacie Loftus, and  
Denise Larson (Division of Intramural Research Genetic Disease  
Research Branch National Human Genome Research Institute, National  
Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive  
MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully  
acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

Location/Qualifiers

1..287  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="G270064G22"  
/cell\_type="melanocyte"  
/clone\_lib="RIKEN full-length enriched, melanocyte"

Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carlini, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Akawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

12466851

22354683

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-ree@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Alawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Cells were provided by Drs. William J Pavan, Stacie Loftus, and

Denise Larson (Division of Intramural Research Genetic Disease

Research Branch National Human Genome Research Institute, National

Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive

MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully

acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

Location/Qualifiers

1. .287

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="G27064G22"

/cell\_type="melanocyte"

/clone\_lib="RIKEN full-length enriched, melanocyte"

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 287;

Best Local Similarity 90.0%; Pred. No. 1.6e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 GGACATGCCGGGATGCC 20

63 GGAGCTGCCGGGATGCC 44

RESULT 43

CK337629 288 bp mRNA linear EST 22-DEC-2003  
LOCUS CK337629  
DEFINITION CK337629-3 NIA Mouse Undifferentiated ES Cell cDNA library (Long)

Mus musculus cDNA clone NIA:CO330A03 IMAGE:30007682 3', mRNA

sequence.

Accession CK337629 GI:40293242

Version CK337629.1

Keywords EST.

Source Mus musculus (house mouse)

Organism Mus musculus

Reference Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 288)

Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.

Construction of long-transcript enriched cDNA libraries from

submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL 11544199

MEDLINE 21429098

PUBMED 11544199

COMMENT Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: ccha@igsun.grc.nia.nih.gov

Plate: CO330 row: A column: 03

Seq primer: -21M13 Forward

High quality sequence stop: 288

PolyA=yes.

FEATURES

source

1. .288

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="129/Sv x 129/Sv-CP"

/db\_xref="niaEST:CO330A03-3"

/db\_xref="taxon:10090"

/clone="NIA:CO330A03 IMAGE:30007682"

/tissue\_type="Undifferentiated ES Cell"

/cell\_line="R1 ES cells"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse Undifferentiated ES Cell cDNA

library (Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (<http://igsun.grc.nia.nih.gov/cDNA>). This is a

long-transcript enriched cDNA library (Ref. Genome Res.

11: 1553-1558 (2001). PMID: 11544199). Total RNAs were

obtained from Dr. Kenneth R. Boheler (National Institute

on Aging, USA). ES cells were cultured without feeder

cells in the presence of LIF and BRL-conditioned media.

Double-stranded cDNAs were synthesized with an Oligo(dT)

primer (Invitrogen):

5'-GACCTAGTCTAGATCGGACGGCGCCCTTTT-3' from

14.2 ug of total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated

to lona-linker l1-SalI, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the

cDNAs were amplified by long-range high fidelity PCR using

Ex Taq polymerase (Takara) with a primer Sal4-S. The

products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes

and cloned into SalI/NotI site of pSPORT1 plasmid vector.

The DH10B E. coli host was transformed with the ligation

mixture by the standard chemical method. The average

insert size is about 2.4 kb. The library was constructed

by Yulan Piao (NIA)."

ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 288;

Best Local Similarity 90.0%; Pred. No. 1.6e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



QY 1 GGACATGCCCGGCATGTCC 20  
 DB 227 GGGCTTGCCCGGCATGTCC 246  
 RESULT 44  
 CK337629 288 bp mRNA linear EST 22-DEC-2003  
 LOCUS C0330A03-3 NIA Mouse Undifferentiated ES Cell cDNA Library (long)  
 DEFINITION Mus musculus cDNA clone NIA:C0330A03 IMAGE:30007682 3', mRNA sequence.  
 ACCESSION CK337629  
 VERSION CK337629.1 GI:40293242  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
 1 (bases 1 to 288)  
 Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method  
 Genome Res. 11 (9), 1553-1558 (2001)  
 JOURNAL  
 MEDLINE 21429098  
 PUBMED 11544199  
 COMMENT Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: C0330 row: A column: 03  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 288  
 POLY-A=yes.  
 Location/Qualifiers  
 1..288  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129/Sv x 129/Sv-Cp"  
 /db\_xref="taxon:10090"  
 /db\_xref="taxon:10090"  
 /clone="NIA:C0330A03 IMAGE:30007682"  
 /tissue\_type="Undifferentiated ES Cell"  
 /cell\_line="R1 ES cells"  
 /lab\_host="DH10B"  
 /clone\_1lb="NIA Mouse Undifferentiated ES Cell cDNA Library (long)"  
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Kenneth R. Boheler (National Institute on Aging, USA). ES cells were cultured without feeder cells in the presence of LIF and BRL-conditioned media. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):  
 5'-PGACTAGTCTAGATCGGACGCGCGCCCTTTT-3' from 14.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao (NIA)."  
 FEATURES  
 source

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 7; Length 288;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGACATGCCCGGCATGTCC 20  
 DB 246 GGACATGCCCGGCATGTCC 227  
 RESULT 45  
 CK340780 288 bp mRNA linear EST 22-DEC-2003  
 LOCUS K0109H06-3 NIA Mouse Hematopoietic Stem Cell (lin-/c-Kit+/Sca-1-)  
 DEFINITION cDNA library (long) Mus musculus cDNA clone NIA:K0109H06  
 IMAGE:30040313 3', mRNA sequence.  
 ACCESSION CK340780  
 VERSION CK340780.1 GI:40296393  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
 1 (bases 1 to 288)  
 Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method  
 Genome Res. 11 (9), 1553-1558 (2001)  
 JOURNAL  
 MEDLINE 21429098  
 PUBMED 11544199  
 COMMENT Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: K0109 row: H column: 06  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 288  
 POLY-A=yes.  
 Location/Qualifiers  
 1..288  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6NCR"  
 /db\_xref="taxon:10090"  
 /db\_xref="taxon:10090"  
 /clone="NIA:K0109H06 IMAGE:30040313"  
 /tissue\_type="Hematopoietic Stem Cell"  
 /lin-/c-Kit+/Sca-1-)"  
 /dev\_stage="Age approx.10 weeks old"  
 /lab\_host="DH10B"  
 /clone\_1lb="NIA Mouse Hematopoietic Stem Cell (lin-/c-Kit+/Sca-1-) cDNA Library (long)"  
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):  
 5'-PGACTAGTCTAGATCGGACGCGCGCCCTTTT-3' from 2.4 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 288;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
|||||  
Db 227 GGGCTTGCCCGGCGCATGTCC 246

RESULT 46  
CK340780/c 288 bp mRNA linear EST 22-DEC-2003  
LOCUS K0109H06-3 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit+/Sca-1-)  
DEFINITION CDNA library (Long) Mus musculus CDNA clone NIA:K0109H06  
IMAGE:30040313 3', mRNA sequence.

ACCESSION CK340780 GI:40296393  
VERSION CK340780.1  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Piao, Y., Ko, N.-T., Lim, M.-K. and Ko, M.-S.H. 1 (bases 1 to 288)  
AUTHORS Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method  
TITLE Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL MEDLINE 21429098  
PUBMED 11544199  
COMMENT Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6620, USA  
Email: cdna@igsun.grc.nia.nih.gov  
Plate: K0109 row: H column: 06  
Seq primer: -21M13 Forward  
High quality sequence stop: 288  
POLY-A-Yes

## FEATURES

## SOURCE

1..288  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6NCR"  
/db\_xref="taxon:10090"  
/db\_xref="taxon:10090"  
/clone="NIA:K0109H06 IMAGE:30040313"  
/tissue\_type="Hematopoietic Stem Cell  
(Lin-/c-Kit+/Sca-1-)"  
/dev\_stage="Age approx.10 weeks old"  
/lab\_host="MDH10B"  
/clone\_lib="NIA Mouse Hematopoietic Stem Cell  
(Lin-/c-Kit+/Sca-1-) CDNA library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse CDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://igsun.grc.nia.nih.gov/CDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were  
obtained from Drs. Dennis Taub, Dan Longo (National  
Institute on Aging, USA), Jonathan Keller (National Cancer  
Institute, USA). Double-stranded cDNAs were synthesized  
with an oligo(dT) primer [Invitrogen:  
5'-pGACTAGTCTAGATCGGACGGCGCCCTTTT-3'] from  
2.4 ug of total RNA, treated with T4 DNA polymerase, and

## ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 288;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20  
|||||  
Db 246 GGACATGCCCGGCGCATGTCC 227

RESULT 47  
BB060455 290 bp mRNA linear EST 25-JUN-2000  
LOCUS BB060455 RIKEN full-length enriched, in vitro fertilized eggs Mus  
DEFINITION musculus CDNA clone 7420450C07 3', mRNA sequence.  
ACCESSION BB060455 GI:8467603  
VERSION BB060455.1  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Kanno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toyota, T., Tsunoda, Y., Wachihi, A., Watanabe, S., Yamamura, T., Yamanka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Kanno, H., et al.)  
JOURNAL Unpublished (2000)  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermolabilization and thermostabilization of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,  
Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
13-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for

# FEATURES

Location/Qualifiers

1..290  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="7420450C07"  
/sex="female"  
/tissue\_type="in vitro fertilized eggs"  
/dev\_stage="egg"  
/lab\_host="DH10B"  
/clone\_1lb="RIKEN full-length enriched, in vitro fertilized eggs"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATTCGAGTTATTAATTAATCCGCCGCCGCC 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTCGAGTTATTAATTAATCCGCCGCCGCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid vector KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 290;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGATGTC 20  
233 GGACATGCCCGGCGAGGCC 252

RESULT 48  
BB060455/c 290 bp mRNA linear EST 25-JUN-2000  
DEFINITION BB060455 RIKEN full-length enriched, in vitro fertilized eggs Mus  
musculus cDNA clone 7420450C07 3', mRNA sequence.  
ACCESSION BB060455  
VERSION BB060455.1 GI:8467603  
KEYWORDS EST.

## ORGANISM

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 290)

## REFERENCE

AUTHORS Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,  
Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
Hirozane, T., Hori, F., Ishi, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,  
Kusakabe, M., Matsumura, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,  
Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,  
Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T.,  
Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,  
Takahashi, P., Tomimaga, N., Toya, T., Tsunoda, Y., Watanabe, A.,  
Watanabe, S., Yamamura, T., Yamahata, I., Yano, R., Yasunishi, A.,  
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.

RIKEN Mouse ESTs (Komno, H., et al.)  
Unpublished (2000)

## TITLE

JOURNAL Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)

# FEATURES

source

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagata, S.,  
Saeki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermolabile and thermolabile activation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Kikuchi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Oza, Y., Muramatsu, M.,  
Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome.ritc.riken.go.jp) for  
further details.

## ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 290;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGATGTC 20  
252 GGCTTGCCTCCGGCGATGTC 233

RESULT 49  
AA289954 293 bp mRNA linear EST 14-APR-1997  
DEFINITION v048b08.r1 Soares mouse lymph node NMIML Mus musculus cDNA clone  
AA289954  
ACCESSION AA289954  
VERSION AA289954.1 GI:1936162  
KEYWORDS EST.

## ORGANISM

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 293)

## REFERENCE

AUTHORS Marra, M., Hallier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Hayashizaki, Y., Itoh, M., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,  
Kusakabe, M., Matsumura, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,  
Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,  
Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T.,  
Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,  
Takahashi, P., Tomimaga, N., Toya, T., Tsunoda, Y., Watanabe, A.,  
Watanabe, S., Yamamura, T., Yamahata, I., Yano, R., Yasunishi, A.,  
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 16:14:28 ; Search time 313.579 Seconds  
(without alignments)  
329.824 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18

Sequence: 1 CGACTGTGAATCCTCCAT 18

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	100.0	1627	9	US-09-917-800A-1491
C 2	18	100.0	1627	16	US-10-368-934-568
C 3	16.4	91.1	390	9	US-09-974-300-3673
C 4	16.4	91.1	724	18	US-10-425-115-32562
C 5	16.4	91.1	1173	14	US-10-038-010-5
C 6	16.4	91.1	1173	15	US-10-267-384-214
C 7	16.4	91.1	1173	18	US-10-746-558-1
C 8	16.4	91.1	1208	15	US-10-392-113-32
C 9	15.4	85.6	441	16	US-10-424-599-102369
C 10	15.4	85.6	458	13	US-10-027-632-285372
C 11	15.4	85.6	458	15	US-10-027-632-285372
C 12	15.4	85.6	584	18	US-10-653-047-6563

13	15.4	85.6	604	18	US-10-425-115-57974	Sequence 57974, A
14	15.4	85.6	608	14	US-10-052-283-349	Sequence 349, App
15	15.4	85.6	616	18	US-10-425-115-180675	Sequence 180675, A
16	15.4	85.6	853	18	US-10-425-115-180677	Sequence 180677, A
17	15.4	85.6	881	18	US-10-425-115-178848	Sequence 178848, A
18	15.4	85.6	1017	18	US-10-425-115-180678	Sequence 180678, A
19	15.4	85.6	1510	9	US-09-731-872-8	Sequence 8, Appl1
20	15.4	85.6	1510	10	US-09-876-997-8	Sequence 8, Appl1
21	15.4	85.6	1523	9	US-09-731-872-47	Sequence 47, Appl1
22	15.4	85.6	1523	10	US-09-876-997-47	Sequence 47, Appl1
23	15.4	85.6	1535	9	US-09-935-390A-15	Sequence 15, Appl1
24	15.4	85.6	1755	9	US-09-745-763-188	Sequence 188, App
25	15.4	85.6	1892	9	US-09-925-298-221	Sequence 221, App
26	15.4	85.6	1892	14	US-10-102-806-221	Sequence 221, App
27	15.4	85.6	154504	17	US-10-322-696-67	Sequence 67, Appl1
28	15.4	85.6	200000	17	US-10-672-764A-33	Sequence 33, Appl1
29	15	83.3	18	14	US-10-131-786-2	Sequence 2, Appl1
30	15	83.3	1024	13	US-10-202-193-5	Sequence 5, Appl1
31	15	83.3	430442	18	US-10-417-375-128	Sequence 128, App
32	14.8	82.2	225	18	US-10-425-115-65084	Sequence 65084, A
33	14.8	82.2	281	11	US-09-922-293-137650	Sequence 137650, A
34	14.8	82.2	350	16	US-10-425-115-47892	Sequence 137650, A
35	14.8	82.2	350	16	US-10-425-115-47892	Sequence 137650, A
36	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
37	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
38	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
39	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
40	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
41	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
42	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
43	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
44	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
45	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
46	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
47	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
48	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
49	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
50	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
51	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
52	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
53	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
54	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
55	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
56	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
57	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
58	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
59	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
60	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
61	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
62	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
63	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
64	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
65	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
66	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
67	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
68	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
69	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
70	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
71	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
72	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
73	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
74	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
75	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
76	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
77	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
78	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
79	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
80	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
81	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
82	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
83	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
84	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1
85	14.4	80.0	19	9	US-10-419-723-1	Sequence 1, Appl1

C 86	14.4	80.0	622	16	US-10-424-599-88609	Sequence 88609, A	C 159	13.8	76.7	50	18	US-10-667-141-28	Sequence 28, Appl
C 87	14.4	80.0	684	13	US-10-027-632-25713	Sequence 25713, A	C 160	13.8	76.7	60	10	US-09-808-975-17568	Sequence 17568, A
C 88	14.4	80.0	684	13	US-10-027-632-25713	Sequence 25713, A	C 161	13.8	76.7	200	9	US-09-864-761-22218	Sequence 22218, A
C 89	14.4	80.0	695	13	US-10-027-632-162409	Sequence 162409, A	C 162	13.8	76.7	201	17	US-10-741-601-25267	Sequence 25267, A
C 90	14.4	80.0	695	13	US-10-027-632-162410	Sequence 162410, A	C 163	13.8	76.7	201	18	US-10-719-993-15787	Sequence 15787, A
C 91	14.4	80.0	695	13	US-10-027-632-162409	Sequence 162409, A	C 164	13.8	76.7	201	18	US-10-719-993-25494	Sequence 25494, A
C 92	14.4	80.0	695	15	US-10-027-632-162409	Sequence 162409, A	C 165	13.8	76.7	201	18	US-10-719-993-28365	Sequence 28365, A
C 93	14.4	80.0	695	15	US-10-027-632-162410	Sequence 162410, A	C 166	13.8	76.7	201	18	US-10-719-993-38413	Sequence 38413, A
C 94	14.4	80.0	864	11	US-09-938-842A-2474	Sequence 2474, Ap	C 167	13.8	76.7	201	18	US-10-719-993-38416	Sequence 38416, A
C 95	14.4	80.0	864	16	US-10-412-699B-219	Sequence 219, App	C 168	13.8	76.7	202	18	US-10-357-930-3881	Sequence 3881, Ap
C 96	14.4	80.0	920	9	US-09-770-445-419	Sequence 419, App	C 169	13.8	76.7	246	9	US-09-974-300-7851	Sequence 7851, Ap
C 97	14.4	80.0	993	17	US-10-437-963-85995	Sequence 85995, A	C 170	13.8	76.7	285	16	US-10-424-599-19280	Sequence 19280, A
C 98	14.4	80.0	1026	16	US-10-425-114-13022	Sequence 13022, A	C 171	13.8	76.7	300	9	US-09-294-093B-4843	Sequence 4843, Ap
C 99	14.4	80.0	1093	9	US-09-186-276B-53	Sequence 53, Appl	C 172	13.8	76.7	312	16	US-10-424-599-43175	Sequence 43175, A
C 100	14.4	80.0	1093	14	US-10-253-007-53	Sequence 53, Appl	C 173	13.8	76.7	320	15	US-10-264-283-47	Sequence 47, Appl
C 101	14.4	80.0	1093	14	US-10-253-007-53	Sequence 53, Appl	C 174	13.8	76.7	341	18	US-10-674-124A-5391	Sequence 5391, Ap
C 102	14.4	80.0	1098	9	US-09-738-626-4368	Sequence 88465, A	C 175	13.8	76.7	345	16	US-10-381-813-113	Sequence 113, App
C 103	14.4	80.0	1541	18	US-10-425-115-88465	Sequence 88465, A	C 176	13.8	76.7	345	16	US-10-381-813-115	Sequence 115, App
C 104	14.4	80.0	1541	18	US-10-425-115-88465	Sequence 88465, A	C 177	13.8	76.7	348	18	US-10-425-115-37196	Sequence 37196, A
C 105	14.4	80.0	1850	17	US-10-433-485A-14	Sequence 32, Appl	C 178	13.8	76.7	375	16	US-10-242-535A-22844	Sequence 22844, A
C 106	14.4	80.0	1850	18	US-10-855-595-3	Sequence 3, Appl1	C 179	13.8	76.7	375	16	US-10-085-783A-22844	Sequence 22844, A
C 107	14.4	80.0	1850	18	US-10-855-532-3	Sequence 3, Appl1	C 180	13.8	76.7	399	16	US-10-424-599-55238	Sequence 55238, A
C 108	14.4	80.0	2574	15	US-10-085-117-243	Sequence 243, App	C 181	13.8	76.7	394	18	US-10-425-115-164935	Sequence 164935, A
C 109	14.4	80.0	2774	15	US-10-278-536-219	Sequence 219, App	C 182	13.8	76.7	400	17	US-10-767-795-2126	Sequence 2126, Ap
C 110	14.4	80.0	2822	18	US-10-723-860-7715	Sequence 7715, Ap	C 183	13.8	76.7	405	18	US-10-357-930-34198	Sequence 34198, A
C 111	14.4	80.0	3417	15	US-10-104-047-808	Sequence 808, App	C 184	13.8	76.7	405	18	US-10-357-930-43060	Sequence 43060, A
C 112	14.4	80.0	4043	9	US-09-989-970-28	Sequence 28, Appl	C 185	13.8	76.7	419	18	US-10-357-930-7901	Sequence 7901, Ap
C 113	14.4	80.0	4967	9	US-09-834-975-558	Sequence 958, App	C 186	13.8	76.7	440	10	US-09-918-993-35436	Sequence 35436, A
C 114	14.4	80.0	4967	10	US-09-814-953-21662	Sequence 21662, A	C 187	13.8	76.7	442	10	US-09-918-993-37849	Sequence 37849, A
C 115	14.4	80.0	5026	15	US-10-085-117-242	Sequence 814, App	C 188	13.8	76.7	443	18	US-10-357-930-11050	Sequence 13050, A
C 116	14.4	80.0	8674	14	US-10-044-090-814	Sequence 51, Appl	C 189	13.8	76.7	447	16	US-10-424-599-136480	Sequence 136480, A
C 117	14.4	80.0	8878	14	US-10-076-016-51	Sequence 5, Appl1	C 190	13.8	76.7	452	9	US-09-867-701-4730	Sequence 4730, Ap
C 118	14.4	80.0	8878	14	US-10-017-724-5	Sequence 512, Ap	C 191	13.8	76.7	459	16	US-10-282-122A-22918	Sequence 22918, A
C 119	14.4	80.0	32195	9	US-09-764-847-1512	Sequence 1512, Ap	C 192	13.8	76.7	459	16	US-09-960-352-1254	Sequence 1254, Ap
C 120	14.4	80.0	32195	14	US-10-092-154-1512	Sequence 7009, App	C 193	13.8	76.7	459	17	US-10-437-963-63280	Sequence 63280, A
C 121	14.4	80.0	42385	18	US-10-719-993-7009	Sequence 676, App	C 194	13.8	76.7	474	9	US-09-864-761-10744	Sequence 10744, Ap
C 122	14.4	80.0	114615	13	US-10-087-192-676	Sequence 241, App	C 195	13.8	76.7	474	9	US-09-864-761-10744	Sequence 10744, Ap
C 123	14.4	80.0	165961	13	US-10-085-117-241	Sequence 2014, Ap	C 196	13.8	76.7	475	18	US-10-357-930-37852	Sequence 37852, A
C 124	14.4	80.0	248436	15	US-10-087-192-2014	Sequence 7065, Ap	C 197	13.8	76.7	486	15	US-10-091-090-225	Sequence 225, App
C 125	14.4	80.0	283351	18	US-10-719-993-7065	Sequence 1, Appl1	C 198	13.8	76.7	486	15	US-09-918-993-20859	Sequence 20859, A
C 126	14.4	80.0	344805	18	US-10-719-993-7065	Sequence 1, Appl1	C 199	13.8	76.7	498	17	US-10-437-963-84363	Sequence 84363, A
C 127	14.4	80.0	1880090	18	US-10-719-993-6615	Sequence 6815, Ap	C 200	13.8	76.7	499	18	US-10-425-115-169404	Sequence 169404, A
C 128	14.4	80.0	3309400	9	US-09-738-626-1	Sequence 1, Appl1	C 201	13.8	76.7	501	9	US-09-864-761-12420	Sequence 12420, A
C 129	14.4	77.8	413	16	US-10-424-599-104931	Sequence 104931, A	C 202	13.8	76.7	504	16	US-10-424-599-64067	Sequence 64067, A
C 130	14.4	77.8	438	13	US-10-027-632-80814	Sequence 80814, A	C 203	13.8	76.7	511	16	US-10-425-114-25440	Sequence 25440, A
C 131	14.4	77.8	438	13	US-10-027-632-80815	Sequence 80815, A	C 204	13.8	76.7	514	9	US-09-604-287A-417	Sequence 417, App
C 132	14.4	77.8	438	15	US-10-027-632-80814	Sequence 80814, A	C 205	13.8	76.7	514	9	US-09-604-287A-417	Sequence 417, App
C 133	14.4	77.8	438	15	US-10-027-632-80815	Sequence 80815, A	C 206	13.8	76.7	514	13	US-09-551-621-417	Sequence 417, App
C 134	14.4	77.8	514	15	US-10-027-632-79374	Sequence 79374, A	C 207	13.8	76.7	514	13	US-10-007-805-417	Sequence 417, App
C 135	14.4	77.8	514	15	US-10-027-632-79374	Sequence 79374, A	C 208	13.8	76.7	514	14	US-10-076-622-417	Sequence 417, App
C 136	14.4	77.8	542	13	US-10-027-632-84022	Sequence 84022, A	C 209	13.8	76.7	514	15	US-10-124-805-417	Sequence 805, A
C 137	14.4	77.8	542	15	US-10-027-632-84022	Sequence 84022, A	C 210	13.8	76.7	516	18	US-10-425-115-178845	Sequence 178845, A
C 138	14.4	77.8	654	13	US-10-027-632-226492	Sequence 226492, A	C 211	13.8	76.7	519	9	US-09-604-287A-258	Sequence 258, App
C 139	14.4	77.8	654	13	US-10-027-632-226493	Sequence 226493, A	C 212	13.8	76.7	519	9	US-09-604-287A-258	Sequence 258, App
C 140	14.4	77.8	654	15	US-10-027-632-226492	Sequence 226492, A	C 213	13.8	76.7	519	9	US-09-339-338-258	Sequence 258, App
C 141	14.4	77.8	654	15	US-10-027-632-226493	Sequence 226493, A	C 214	13.8	76.7	519	10	US-09-551-621-258	Sequence 258, App
C 142	14.4	77.8	733	13	US-10-027-632-21556	Sequence 21556, A	C 215	13.8	76.7	519	13	US-10-007-805-258	Sequence 258, App
C 143	14.4	77.8	733	13	US-10-027-632-21556	Sequence 21556, A	C 216	13.8	76.7	519	13	US-10-076-622-258	Sequence 258, App
C 144	14.4	77.8	940	16	US-10-424-599-77765	Sequence 77765, A	C 217	13.8	76.7	519	15	US-10-124-805-258	Sequence 258, App
C 145	14.4	77.8	1234	9	US-09-778-844-86	Sequence 86, Appl1	C 218	13.8	76.7	529	18	US-10-357-930-15927	Sequence 15927, A
C 146	14.4	77.8	1395	15	US-10-369-493-25339	Sequence 25439, A	C 219	13.8	76.7	539	18	US-10-425-115-132876	Sequence 132876, A
C 147	14.4	77.8	1395	16	US-10-793-639-167	Sequence 167, App	C 220	13.8	76.7	548	18	US-10-425-115-159293	Sequence 159293, A
C 148	14.4	77.8	3438	16	US-10-062-674-1639	Sequence 1639, Ap	C 221	13.8	76.7	551	16	US-10-240-425-152	Sequence 152, App
C 149	14.4	77.8	13524	18	US-10-719-993-6913	Sequence 6913, Ap	C 222	13.8	76.7	558	16	US-10-425-114-21890	Sequence 21890, App
C 150	14.4	77.8	13904	15	US-10-017-161-1977	Sequence 1977, Ap	C 223	13.8	76.7	561	16	US-10-425-114-22055	Sequence 22055, A
C 151	14.4	77.8	13904	15	US-10-292-798-1625	Sequence 1625, Ap	C 224	13.8	76.7	564	18	US-10-425-115-169406	Sequence 169406, A
C 152	14.4	77.8	39412	17	US-10-043-160-31	Sequence 31, Appl	C 225	13.8	76.7	565	18	US-10-357-930-45755	Sequence 45755, A
C 153	14.4	77.8	255439	18	US-10-043-160-5	Sequence 5, Appl1	C 226	13.8	76.7	567	16	US-10-424-599-90678	Sequence 90678, A
C 154	14.4	77.8	255439	18	US-10-719-993-6799	Sequence 6799, Ap	C 227	13.8	76.7	568	18	US-10-793-032-4	Sequence 4, Appl1
C 155	13.8	76.7	25	15	US-10-098-263B-115019	Sequence 115019, A	C 228	13.8	76.7	572	13	US-10-027-632-225533	Sequence 225533, A
C 156	13.8	76.7	34	10	US-09-747-377-309	Sequence 309, App	C 229	13.8	76.7	572	13	US-10-027-632-225533	Sequence 225533, A
C 157	13.8	76.7	34	10	US-10-105-613-309	Sequence 309, App	C 230	13.8	76.7	575	13	US-10-027-632-128672	Sequence 128672, A
C 158	13.8	76.7	49	18	US-10-667-141-27	Sequence 27, Appl	C 231	13.8	76.7	575	15	US-10-027-632-128672	Sequence 128672, A

232	13.8	76.7	576	16	US-10-425-114-77205	Sequence 27205, A	C 305	13.8	76.7	967	15	US-10-106-698-760	Sequence 760, App
C 233	13.8	76.7	582	17	US-10-021-323-7768	Sequence 7768, App	C 306	13.8	76.7	981	16	US-10-424-599-74187	Sequence 74187, A
C 234	13.8	76.7	589	18	US-10-357-930-56981	Sequence 56981, A	C 307	13.8	76.7	990	17	US-10-437-963-99283	Sequence 99283, A
C 235	13.8	76.7	598	13	US-10-027-632-230606	Sequence 230606, A	C 308	13.8	76.7	999	16	US-10-425-114-25917	Sequence 25917, A
C 236	13.8	76.7	598	15	US-10-027-632-230606	Sequence 230606, A	C 309	13.8	76.7	1016	9	US-09-823-2454-451	Sequence 451, App
C 237	13.8	76.7	601	13	US-10-027-632-24078	Sequence 24078, A	C 310	13.8	76.7	1018	17	US-10-767-701-5775	Sequence 5775, App
C 238	13.8	76.7	601	13	US-10-027-632-133294	Sequence 133294, A	C 311	13.8	76.7	1021	16	US-10-321-039-17	Sequence 17, App1
C 239	13.8	76.7	601	13	US-10-027-632-133295	Sequence 133295, A	C 312	13.8	76.7	1032	15	US-10-074-511-58	Sequence 58, App1
C 240	13.8	76.7	601	15	US-10-027-632-24078	Sequence 24078, A	C 313	13.8	76.7	1035	16	US-10-425-114-22420	Sequence 22420, A
C 241	13.8	76.7	601	15	US-10-027-632-133294	Sequence 133294, A	C 314	13.8	76.7	1041	15	US-10-401-307-79	Sequence 79, App1
C 242	13.8	76.7	603	17	US-10-027-632-133295	Sequence 133295, A	C 315	13.8	76.7	1046	9	US-09-925-297-295	Sequence 295, App
C 243	13.8	76.7	606	18	US-10-425-115-116979	Sequence 65496, A	C 316	13.8	76.7	1047	16	US-10-425-114-22427	Sequence 22427, A
C 244	13.8	76.7	610	13	US-10-027-632-257353	Sequence 257353, A	C 317	13.8	76.7	1053	9	US-09-770-445-157	Sequence 157, App
C 245	13.8	76.7	610	13	US-10-027-632-257353	Sequence 257353, A	C 318	13.8	76.7	1059	16	US-10-425-114-7086	Sequence 7086, App
C 246	13.8	76.7	610	15	US-10-027-632-257353	Sequence 257353, A	C 319	13.8	76.7	1059	16	US-10-425-114-22421	Sequence 22421, A
C 247	13.8	76.7	612	17	US-10-437-963-12552	Sequence 12592, A	C 320	13.8	76.7	1107	18	US-10-425-115-149921	Sequence 149921, A
C 248	13.8	76.7	621	13	US-10-037-963-84362	Sequence 84362, A	C 321	13.8	76.7	1130	16	US-10-282-1224-20074	Sequence 20074, A
C 249	13.8	76.7	627	15	US-10-027-632-185178	Sequence 185178, A	C 322	13.8	76.7	1238	18	US-10-425-115-42420	Sequence 42420, A
C 250	13.8	76.7	627	15	US-10-027-632-185178	Sequence 185178, A	C 323	13.8	76.7	1238	18	US-10-425-115-42420	Sequence 42420, A
C 251	13.8	76.7	642	13	US-10-027-632-48134	Sequence 48134, A	C 324	13.8	76.7	1272	16	US-10-282-1224-35694	Sequence 35694, A
C 252	13.8	76.7	642	13	US-10-027-632-48134	Sequence 48134, A	C 325	13.8	76.7	1302	16	US-10-424-599-17155	Sequence 17155, A
C 253	13.8	76.7	647	18	US-10-425-115-178847	Sequence 178847, A	C 326	13.8	76.7	1305	16	US-10-282-1224-20074	Sequence 20074, A
C 254	13.8	76.7	652	13	US-10-027-632-105460	Sequence 105460, A	C 327	13.8	76.7	1305	17	US-10-437-963-19573	Sequence 19573, A
C 255	13.8	76.7	652	15	US-10-027-632-105460	Sequence 105460, A	C 328	13.8	76.7	1316	18	US-10-425-115-135105	Sequence 135105, A
C 256	13.8	76.7	661	17	US-10-437-963-100249	Sequence 100249, A	C 329	13.8	76.7	1373	9	US-09-263-959-285	Sequence 285, App
C 257	13.8	76.7	677	13	US-10-027-632-243275	Sequence 243275, A	C 330	13.8	76.7	1395	9	US-09-938-8424-724	Sequence 724, App
C 258	13.8	76.7	677	15	US-10-027-632-243275	Sequence 243275, A	C 331	13.8	76.7	1395	11	US-09-938-8424-724	Sequence 724, App
C 259	13.8	76.7	682	13	US-10-027-632-22630	Sequence 22630, A	C 332	13.8	76.7	1431	10	US-09-992-6008-87	Sequence 87, App1
C 260	13.8	76.7	682	15	US-10-027-632-22630	Sequence 22630, A	C 333	13.8	76.7	1431	10	US-09-992-6008-87	Sequence 87, App1
C 261	13.8	76.7	684	16	US-10-424-599-135530	Sequence 135530, A	C 334	13.8	76.7	1431	10	US-09-924-140-89	Sequence 89, App1
C 262	13.8	76.7	687	18	US-10-357-930-22538	Sequence 22538, A	C 335	13.8	76.7	1431	10	US-09-924-140-89	Sequence 89, App1
C 263	13.8	76.7	687	18	US-10-357-930-22538	Sequence 22538, A	C 336	13.8	76.7	1431	10	US-09-992-0958-87	Sequence 87, App1
C 264	13.8	76.7	696	18	US-10-425-115-148874	Sequence 148874, A	C 337	13.8	76.7	1431	10	US-09-992-0958-87	Sequence 87, App1
C 265	13.8	76.7	700	9	US-09-920-3004-1683	Sequence 1683, App	C 338	13.8	76.7	1431	10	US-09-992-0958-87	Sequence 87, App1
C 266	13.8	76.7	700	15	US-10-033-528-1683	Sequence 1683, App	C 339	13.8	76.7	1431	14	US-10-000-489-87	Sequence 87, App1
C 267	13.8	76.7	708	18	US-10-099-926-1683	Sequence 1683, App	C 340	13.8	76.7	1431	14	US-10-000-489-87	Sequence 87, App1
C 268	13.8	76.7	708	18	US-10-099-926-1683	Sequence 1683, App	C 341	13.8	76.7	1431	14	US-10-000-489-87	Sequence 87, App1
C 269	13.8	76.7	709	16	US-10-424-599-69223	Sequence 69223, A	C 342	13.8	76.7	1431	14	US-10-000-986-87	Sequence 87, App1
C 270	13.8	76.7	714	13	US-10-027-632-16482	Sequence 16482, A	C 343	13.8	76.7	1431	14	US-10-000-986-87	Sequence 87, App1
C 271	13.8	76.7	714	15	US-10-027-632-16482	Sequence 16482, A	C 344	13.8	76.7	1431	15	US-10-154-678-89	Sequence 89, App1
C 272	13.8	76.7	735	17	US-10-767-701-14807	Sequence 14807, A	C 345	13.8	76.7	1431	15	US-10-154-678-89	Sequence 89, App1
C 273	13.8	76.7	735	13	US-10-027-632-135508	Sequence 135508, A	C 346	13.8	76.7	1431	15	US-10-001-142-87	Sequence 87, App1
C 274	13.8	76.7	735	15	US-10-027-632-135508	Sequence 135508, A	C 347	13.8	76.7	1431	15	US-10-001-142-87	Sequence 87, App1
C 275	13.8	76.7	737	18	US-10-425-115-172628	Sequence 172628, A	C 348	13.8	76.7	1433	18	US-10-425-115-52898	Sequence 52898, A
C 276	13.8	76.7	739	13	US-10-027-632-101426	Sequence 101426, A	C 349	13.8	76.7	1506	9	US-09-886-241-4	Sequence 41, App1
C 277	13.8	76.7	743	15	US-10-027-632-101426	Sequence 101426, A	C 350	13.8	76.7	1513	9	US-09-920-3004-1682	Sequence 1682, App
C 278	13.8	76.7	748	17	US-10-437-963-59595	Sequence 59595, A	C 351	13.8	76.7	1513	13	US-10-033-528-1682	Sequence 1682, App
C 279	13.8	76.7	754	18	US-10-425-115-64349	Sequence 64349, A	C 352	13.8	76.7	1513	15	US-10-205-823-116	Sequence 116, App
C 280	13.8	76.7	754	18	US-10-425-115-64349	Sequence 64349, A	C 353	13.8	76.7	1513	15	US-10-177-993-153	Sequence 153, App
C 281	13.8	76.7	758	13	US-10-027-632-110325	Sequence 110325, A	C 354	13.8	76.7	1513	15	US-10-099-926-1682	Sequence 1682, App
C 282	13.8	76.7	758	15	US-10-027-632-110325	Sequence 110325, A	C 355	13.8	76.7	1518	17	US-10-738-455-2	Sequence 2, App1
C 283	13.8	76.7	759	13	US-10-027-632-29911	Sequence 29911, A	C 356	13.8	76.7	1527	18	US-10-425-115-150275	Sequence 150275, A
C 284	13.8	76.7	765	9	US-09-912-787-11	Sequence 29911, A	C 357	13.8	76.7	1551	17	US-10-437-963-439	Sequence 439, App
C 285	13.8	76.7	765	9	US-09-912-787-11	Sequence 29911, A	C 358	13.8	76.7	1567	18	US-10-425-115-9029	Sequence 9029, App
C 286	13.8	76.7	765	9	US-09-912-787-11	Sequence 29911, A	C 359	13.8	76.7	1668	18	US-10-739-830-439	Sequence 439, App
C 287	13.8	76.7	765	13	US-10-014-326-56	Sequence 56, App1	C 360	13.8	76.7	1677	18	US-10-044-090-140	Sequence 140, App
C 288	13.8	76.7	765	13	US-10-014-326-56	Sequence 56, App1	C 361	13.8	76.7	1718	18	US-10-425-115-29698	Sequence 29698, A
C 289	13.8	76.7	766	18	US-10-425-115-117779	Sequence 117779, A	C 362	13.8	76.7	1781	13	US-10-027-632-200093	Sequence 200093, A
C 290	13.8	76.7	781	18	US-10-425-115-114438	Sequence 114438, A	C 363	13.8	76.7	1781	13	US-10-027-632-200093	Sequence 200093, A
C 291	13.8	76.7	782	17	US-10-767-701-14806	Sequence 14806, A	C 364	13.8	76.7	1800	18	US-10-739-930-89	Sequence 89, App1
C 292	13.8	76.7	792	9	US-09-938-8424-1950	Sequence 1950, App	C 365	13.8	76.7	1841	16	US-10-435-696-66	Sequence 66, App1
C 293	13.8	76.7	792	11	US-09-938-8424-1950	Sequence 1950, App	C 366	13.8	76.7	1853	9	US-09-745-783-84	Sequence 84, App1
C 294	13.8	76.7	792	16	US-10-425-114-35743	Sequence 35743, A	C 367	13.8	76.7	1893	16	US-10-424-599-42110	Sequence 42110, A
C 295	13.8	76.7	859	9	US-09-764-877-971	Sequence 971, App	C 368	13.8	76.7	2002	9	US-09-925-101-330	Sequence 230, App
C 296	13.8	76.7	859	16	US-10-242-515-971	Sequence 971, App	C 369	13.8	76.7	2020	9	US-09-984-245-49	Sequence 49, App1
C 297	13.8	76.7	880	18	US-10-425-115-135103	Sequence 135103, App	C 370	13.8	76.7	2020	10	US-09-966-262-49	Sequence 49, App1
C 298	13.8	76.7	887	18	US-10-425-115-147019	Sequence 147019, A	C 371	13.8	76.7	2020	10	US-09-983-866-49	Sequence 49, App1
C 299	13.8	76.7	891	17	US-10-437-963-94739	Sequence 94739, A	C 372	13.8	76.7	2020	14	US-10-059-395-49	Sequence 49, App1
C 300	13.8	76.7	900	17	US-10-437-963-94628	Sequence 94628, A	C 373	13.8	76.7	2020	14	US-10-143-090-49	Sequence 49, App1
C 301	13.8	76.7	905	9	US-09-954-106-1198	Sequence 1198, App	C 374	13.8	76.7	2067	16	US-10-282-1224-26924	Sequence 26924, A
C 302	13.8	76.7	905	9	US-09-880-107-424	Sequence 424, App	C 375	13.8	76.7	2129	17	US-10-437-963-26460	Sequence 26460, A
C 303	13.8	76.7	915	15	US-10-369-493-33793	Sequence 33793, A	C 376	13.8	76.7	2179	17	US-10-115-635-117	Sequence 117, App
C 304	13.8	76.7	946	18	US-10-425-115-147008	Sequence 147008, A	C 377	13.8	76.7	2199	15	US-10-205-194-56	Sequence 56, App1

C 378	13.8	76.7	2199	16	US-10-388-934-636	Sequence 636, App
379	13.8	76.7	2238	15	US-10-260-238-529	Sequence 529, App
380	13.8	76.7	2244	17	US-10-437-963-13344	Sequence 13344, A
381	13.8	76.7	2207	15	US-10-120-968-114	Sequence 314, App
C 382	13.8	76.7	2340	17	US-10-437-963-65515	Sequence 65515, A
383	13.8	76.7	2420	18	US-10-357-930-25066	Sequence 25066, A
384	13.8	76.7	2455	16	US-10-424-599-124227	Sequence 124227, A
C 385	13.8	76.7	2510	16	US-10-108-260A-1354	Sequence 1354, App
C 386	13.8	76.7	2518	13	US-10-027-632-103157	Sequence 103157, A
C 387	13.8	76.7	2518	13	US-10-027-632-103158	Sequence 103158, A
C 388	13.8	76.7	2518	13	US-10-027-632-103157	Sequence 103157, A
C 389	13.8	76.7	2518	13	US-10-027-632-103158	Sequence 103158, A
C 390	13.8	76.7	2568	15	US-10-128-714-488	Sequence 488, App
C 391	13.8	76.7	2683	15	US-10-128-714-5488	Sequence 5488, App
C 392	13.8	76.7	2718	9	US-09-764-877-3940	Sequence 3940, App
C 393	13.8	76.7	2718	16	US-10-242-515-9940	Sequence 3940, App
C 394	13.8	76.7	3166	16	US-10-108-260A-739	Sequence 739, App
C 395	13.8	76.7	3194	18	US-10-723-860-7737	Sequence 7737, App
C 396	13.8	76.7	3216	17	US-10-437-963-74161	Sequence 74161, A
C 397	13.8	76.7	3858	10	US-09-814-353-21751	Sequence 21751, A
C 398	13.8	76.7	4161	14	US-10-198-846-13097	Sequence 13097, A
C 399	13.8	76.7	4406	17	US-10-437-963-13891	Sequence 13891, A
C 400	13.8	76.7	4518	18	US-10-483-512-73	Sequence 73, App
C 401	13.8	76.7	6393	15	US-10-116-802-123	Sequence 123, App
C 402	13.8	76.7	6393	15	US-10-334-726-323	Sequence 323, App
C 403	13.8	76.7	6394	15	US-10-177-293-363	Sequence 363, App
C 404	13.8	76.7	6415	15	US-10-172-118-1542	Sequence 1542, App
C 405	13.8	76.7	6415	16	US-10-342-887-1542	Sequence 1542, App
C 406	13.8	76.7	6641	18	US-10-723-860-7301	Sequence 7301, App
C 407	13.8	76.7	6930	10	US-09-770-107-1	Sequence 1, App
C 408	13.8	76.7	7450	18	US-10-788-792-11	Sequence 11, App
C 409	13.8	76.7	7450	18	US-10-723-860-1915	Sequence 1915, App
C 410	13.8	76.7	7473	14	US-10-723-860-6240	Sequence 6240, App
C 411	13.8	76.7	8818	18	US-10-175-523-63	Sequence 63, App
C 412	13.8	76.7	9903	10	US-09-764-891-8188	Sequence 8188, App
C 413	13.8	76.7	10785	9	US-09-764-878-324	Sequence 324, App
C 414	13.8	76.7	10785	14	US-10-079-854-324	Sequence 328, App
C 415	13.8	76.7	13315	9	US-09-764-878-328	Sequence 328, App
C 416	13.8	76.7	13315	14	US-10-079-854-328	Sequence 328, App
C 417	13.8	76.7	15654	18	US-10-667-141-58	Sequence 58, App
C 418	13.8	76.7	15654	18	US-10-667-141-59	Sequence 59, App
C 419	13.8	76.7	15654	18	US-10-667-141-60	Sequence 60, App
C 420	13.8	76.7	20261	9	US-09-764-878-325	Sequence 325, App
C 421	13.8	76.7	20261	14	US-10-079-854-325	Sequence 325, App
C 422	13.8	76.7	29328	17	US-10-367-094-189	Sequence 189, App
C 423	13.8	76.7	32835	11	US-09-997-722-214	Sequence 214, App
C 424	13.8	76.7	41079	17	US-10-322-281-516	Sequence 516, App
C 425	13.8	76.7	51001	16	US-10-189-268-11	Sequence 11, App
C 426	13.8	76.7	55966	11	US-09-997-722-7	Sequence 7, App
C 427	13.8	76.7	57013	13	US-10-087-192-1798	Sequence 1798, App
C 428	13.8	76.7	62822	18	US-10-087-192-1753	Sequence 6881, App
C 429	13.8	76.7	67076	13	US-10-087-192-1753	Sequence 1753, App
C 430	13.8	76.7	74580	18	US-10-719-993-6830	Sequence 6830, App
C 431	13.8	76.7	88277	18	US-10-719-993-6931	Sequence 6931, App
C 432	13.8	76.7	89047	16	US-10-672-787-34	Sequence 34, App
C 433	13.8	76.7	95960	13	US-10-087-192-1384	Sequence 1384, App
C 434	13.8	76.7	95960	13	US-10-087-192-1390	Sequence 1390, App
C 435	13.8	76.7	96598	11	US-09-997-722-127	Sequence 127, App
C 436	13.8	76.7	101193	17	US-10-322-281-468	Sequence 468, App
C 437	13.8	76.7	104514	13	US-10-087-192-1391	Sequence 1391, App
C 438	13.8	76.7	122186	10	US-09-563-728A-36	Sequence 36, App
C 439	13.8	76.7	133955	15	US-10-087-192-1244	Sequence 1984, App
C 440	13.8	76.7	136726	15	US-10-085-117-244	Sequence 244, App
C 441	13.8	76.7	154817	17	US-10-085-117-334	Sequence 334, App
C 442	13.8	76.7	168174	15	US-10-071-411-63	Sequence 63, App
C 443	13.8	76.7	168273	17	US-10-071-411-2	Sequence 2, App
C 444	13.8	76.7	200000	17	US-10-672-764A-31	Sequence 31, App
C 445	13.8	76.7	230093	18	US-10-719-993-6861	Sequence 6861, App
C 446	13.8	76.7	230101	18	US-10-719-993-6829	Sequence 6829, App
C 447	13.8	76.7	238910	18	US-10-719-993-6801	Sequence 6801, App
C 448	13.8	76.7	254087	13	US-10-087-192-223	Sequence 223, App
C 449	13.8	76.7	257156	11	US-09-968-007A-427	Sequence 427, App
C 450	13.8	76.7	302603	16	US-10-271-416-8	Sequence 8, App

## ALIGNMENTS

RESULT 1  
 US-09-917-800A-1491/C  
 ; Sequence 1491, Application US/09917800A  
 ; Patent No. US20020119462A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mendrick, Donna  
 ; APPLICANT: Porter, Mark  
 ; APPLICANT: Johnson, Kory  
 ; APPLICANT: Castle, Arthur  
 ; APPLICANT: Elashoff, Michael  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Molecular Toxicology Modeling  
 ; FILE REFERENCE: 44921-5038-US  
 ; CURRENT APPLICATION NUMBER: US/09/917, 800A  
 ; PRIORITY FILING DATE: 2001-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/222,040  
 ; PRIOR FILING DATE: 2000-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/222,880



PRIOR FILING DATE: 2000-11-02  
PRIOR APPLICATION NUMBER: US 60/290,029  
PRIOR FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: US 60/290,645  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: US 60/292,336  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: US 60/295,798  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 60/297,457  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,884  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/303,459  
NUMBER OF SEQ ID NOS: 1740  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 1491  
LENGTH: 1627  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020119462A1 X13058  
US-09-917-800A-1491

Query Match  
Best Local Similarity 100.0%; Score 18; DB 9; Length 1627;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18  
|||||  
DB 41 CGACTGTGATCCTCCAT 24

RESULT 2  
US-10-388-934-568/c  
Sequence 568, Application US/10388934  
Publication No. US20040005547A1  
GENERAL INFORMATION:  
APPLICANT: Boess, Franziska  
APPLICANT: Suter-Dick, Laura  
APPLICANT: Wolf, Detlef  
TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY  
FILE REFERENCE: 21199  
CURRENT APPLICATION NUMBER: US/10/388,934  
CURRENT FILING DATE: 2003-03-14  
PRIOR APPLICATION NUMBER: 02005336.9  
PRIOR FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 02015657.6  
PRIOR FILING DATE: 2002-07-17  
NUMBER OF SEQ ID NOS: 862  
SOFTWARE: Patent In version 3.1  
SEQ ID NO 568  
LENGTH: 1627  
TYPE: DNA  
ORGANISM: Rattus norvegicus (No. US20040005547A1 way rat)  
US-10-388-934-568

Query Match  
Best Local Similarity 100.0%; Score 18; DB 16; Length 1627;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18  
|||||  
DB 41 CGACTGTGATCCTCCAT 24

RESULT 3  
US-09-974-300-3673/c  
Sequence 3673, Application US/09974300  
Patent No. US20020146721A1  
GENERAL INFORMATION:  
APPLICANT: Berka, Randy M.

APPLICANT: Clausen, Ib Groth  
TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
FILE REFERENCE: 10085.500-US  
CURRENT APPLICATION NUMBER: US/09/974,300  
CURRENT FILING DATE: 2001-10-05  
PRIOR APPLICATION NUMBER: 09/680,598  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/279,526  
PRIOR FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 8481  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3673  
LENGTH: 390  
TYPE: DNA  
ORGANISM: Bacillus licheniformis  
US-09-974-300-3673

Query Match  
Best Local Similarity 91.1%; Score 16.4; DB 9; Length 390;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18  
|||||  
DB 120 CGACTGTGATCCTCCAT 103

RESULT 4  
US-10-425-115-32582/c  
Sequence 32582, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 32582  
LENGTH: 724  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1) ..(724)  
OTHER INFORMATION: unsure at all n locations  
FEATURE:  
OTHER INFORMATION: Clone ID: MFT4577\_129729C.1  
US-10-425-115-32582

Query Match  
Best Local Similarity 94.4%; Score 16.4; DB 18; Length 724;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18  
|||||  
DB 176 CGACTGTGATCCTCCAT 159

RESULT 5  
US-10-038-010-5/c  
Sequence 5, Application US/10038010  
Publication No. US20030040089A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS  
APPLICANT: Pierre, Legrain  
TITLE OF INVENTION: Protein-protein interactions in adipocyte cells  
FILE REFERENCE: B4767A  
CURRENT APPLICATION NUMBER: US/10/038,010

CURRENT FILING DATE: 2002-07-23  
PRIOR APPLICATION NUMBER: US 60/259,377  
PRIOR FILING DATE: 2001-01-02  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: mouse p53  
FEATURE:  
NAME/KEY: gene  
LOCATION: (1)..(1173)  
OTHER INFORMATION: mouse p53 : Tumour suppressor protein  
US-10-038-010-5

Query Match 91.1%; Score 16.4; DB 14; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 6  
US-10-267-384-214/c  
Sequence 214, Application US/10267384  
Publication No. US20030198623A1  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
Tartaglia, James  
Cox, William I.  
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 217  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtlis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/267,384  
FILING DATE: 09-Oct-2002  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2530  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 42506CURTWS  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 214:  
US-10-267-384-214

Query Match 91.1%; Score 16.4; DB 15; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCAT 18

Db 27 CGACTGTGAATCCTCCAT 10  
|||||

RESULT 7  
US-10-746-558-1/c  
Sequence 1, Application US/10746558  
Publication No. US20040208650A1  
GENERAL INFORMATION:  
APPLICANT: Ellenhorn, Joshua D.I.  
APPLICANT: Diamond, Don J.  
TITLE OF INVENTION: Modified vaccinia Ankara expressing p53 in cancer immunotherapy  
FILE REFERENCE: 54435,8005 US00  
CURRENT APPLICATION NUMBER: US/10/746,558  
CURRENT FILING DATE: 2003-12-23  
PRIOR APPLICATION NUMBER: 06/436,268  
PRIOR FILING DATE: 2002-12-23  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-746-558-1

Query Match 91.1%; Score 16.4; DB 18; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 8  
US-10-392-113-32/c  
Sequence 32, Application US/10392113  
Publication No. US20030224993A1  
GENERAL INFORMATION:  
APPLICANT: Land, Hartmut  
APPLICANT: Delieu, Laurent  
TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION  
FILE REFERENCE: 21108,0005U3  
CURRENT APPLICATION NUMBER: US/10/392,113  
CURRENT FILING DATE: 2003-03-17  
PRIOR APPLICATION NUMBER: 60/365,078  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: PCT/US01/32127  
PRIOR FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 60/239,705  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 32  
LENGTH: 1208  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:/Note =  
US-10-392-113-32

Query Match 91.1%; Score 16.4; DB 15; Length 1208;  
Best Local Similarity 94.4%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 52 CGACTGTGAATCCTCCAT 35

RESULT 9

US-10-424-599-102369/C  
; Sequence 102369, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 102369  
; LENGTH: 441  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_63456C.1  
US-10-424-599-102369

Query Match 85.6%; Score 15.4; DB 16; Length 441;  
Best Local Similarity 94.1%; Pred. No. 4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18  
DB 236 GAATGTGAATCTCCAT 220

RESULT 10  
US-10-027-632-285372  
; Sequence 285372, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 285372  
; LENGTH: 458  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-285372

Query Match 85.6%; Score 15.4; DB 13; Length 458;  
Best Local Similarity 94.1%; Pred. No. 4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18  
DB 235 GACTGTGAATCTCCAT 251

RESULT 11  
US-10-027-632-285372  
; Sequence 285372, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 285372  
; LENGTH: 458  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-285372

Query Match 85.6%; Score 15.4; DB 15; Length 458;  
Best Local Similarity 94.1%; Pred. No. 4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18  
DB 235 GACTGTGAATCTCCAT 251

RESULT 12  
US-10-653-047-6563  
; Sequence 6563, Application US/10653047  
; Publication No. US20040229367A1  
; GENERAL INFORMATION:  
; APPLICANT: Randy M. Berka  
; APPLICANT: Michael W. Rey  
; APPLICANT: Jeffrey R. Shuster  
; APPLICANT: Sakari Kauppinen  
; APPLICANT: Ib Groth Clausen  
; APPLICANT: Peter Bjørke Olsen  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 5849.200-US  
; CURRENT APPLICATION NUMBER: US/10/653,047  
; CURRENT FILING DATE: 2003-08-29  
; PRIOR APPLICATION NUMBER: US/09/533,559  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/273,623  
; PRIOR FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 7860  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6563  
; LENGTH: 584  
; TYPE: DNA  
; ORGANISM: Aspergillus oryzae  
US-10-653-047-6563

Query Match 85.6%; Score 15.4; DB 18; Length 584;  
Best Local Similarity 94.1%; Pred. No. 4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACTGTGAATCCTCCAT 18  
|||  
DB 240 GATTGTGAATCCTCCAT 256

## RESULT 13

US-10-425-115-57974  
; Sequence 57974, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 57974  
; LENGTH: 604  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(604)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_15286C.1  
US-10-425-115-57974

Query Match 85.6%; Score 15.4; DB 18; Length 604;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCA 17  
|||  
DB 420 CGACTGTGAATCCTCCA 436

## RESULT 14

US-10-052-283-349  
; Sequence 349, Application US/10052283  
; Publication No. US20030064379A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND METHOD OF USE THEREOF  
; FILE REFERENCE: P2751R1C1  
; CURRENT APPLICATION NUMBER: US/10/052,283  
; CURRENT FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: PCT/US00/20006  
; PRIOR FILING DATE: 2000-07-21  
; PRIOR APPLICATION NUMBER: US 60/145,701  
; PRIOR FILING DATE: 1999-07-26  
; NUMBER OF SEQ ID NOS: 564  
; SEQ ID NO 349  
; LENGTH: 608  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 19, 46, 101, 110, 132, 271  
; OTHER INFORMATION: unknown base  
US-10-052-283-349

Query Match 85.6%; Score 15.4; DB 14; Length 608;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCA 17  
|||  
DB 299 CGACTGTGAATCCTCCA 315

## RESULT 15

US-10-425-115-180675  
; Sequence 180675, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 180675  
; LENGTH: 616  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(604)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_96355C.1  
US-10-425-115-180675

Query Match 85.6%; Score 15.4; DB 18; Length 616;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACTGTGAATCCTCCAT 18  
|||  
DB 262 GACTGTGAATCCTCCAT 278

## RESULT 16

US-10-425-115-180677  
; Sequence 180677, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 180677  
; LENGTH: 853  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 19, 46, 101, 110, 132, 271  
; OTHER INFORMATION: unknown base  
US-10-425-115-180677

Query Match 85.6%; Score 15.4; DB 18; Length 853;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACTGTGAATCCTCCAT 18  
|||  
DB 462 GACTGTGAATCCTCCAT 478

## RESULT 17

US-10-425-115-178848  
; Sequence 178848, Application US/10425115  
; Publication No. US20040214272A1

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 178848
; LENGTH: 881
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_94693C.1
; US-10-425-115-178848

Query Match      85.6%; Score 15.4; DB 18; Length 881;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 GACTGTGAATCCTCCAT 18
      |||||
Db      468 GACTGTGAATCCTCCAT 484

RESULT 18
; US-10-425-115-180678
; Sequence 180678, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180678
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1017)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_96358C.1
; US-10-425-115-180678

Query Match      85.6%; Score 15.4; DB 18; Length 1017;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 GACTGTGAATCCTCCAT 18
      |||||
Db      475 GACTGTGAATCCTCCAT 491

RESULT 19
; US-09-731-872-8
; Sequence 8, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
```

```

; FILE REFERENCE: 78-US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 8
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..1398
; NAME/KEY: sig_peptide
; LOCATION: 190..252
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.8172934575094
; OTHER INFORMATION: seq ALLWAQEVGVLA/GR
; US-09-876-997-8

Query Match      85.6%; Score 15.4; DB 10; Length 1510;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 CGACTGTGAATCCTCCA 17
      |||||
Db      405 CGACTGTGAATCCTCCA 421

RESULT 20
; US-09-876-997-8
; Sequence 8, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78-US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 8
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..1398
; NAME/KEY: sig_peptide
; LOCATION: 190..252
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.8172934575094
; OTHER INFORMATION: seq ALLWAQEVGVLA/GR
; US-09-876-997-8

Query Match      85.6%; Score 15.4; DB 10; Length 1510;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 CGACTGTGAATCCTCCA 17
      |||||
Db      405 CGACTGTGAATCCTCCA 421
```

RESULT 21  
US-09-731-872-47  
; Sequence 47, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78-US3-REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 47  
; LENGTH: 1523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 217..1410  
; NAME/KEY: sig\_peptide  
; LOCATION: 217..279  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 5.8172934575094  
; OTHER INFORMATION: seq ALLMAQEVGYLA/GR  
US-09-731-872-47

Query Match 85.6%; Score 15.4; DB 9; Length 1523;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCA 17  
Db 432 CGACTGTGAATCCTCCA 448

RESULT 22  
US-09-876-997-47  
; Sequence 47, Application US/09876997  
; Publication No. US20030152921A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78-US4-CTP  
; CURRENT APPLICATION NUMBER: US/09/876,997  
; PRIOR FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 47  
; LENGTH: 1523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 217..1410  
; NAME/KEY: sig\_peptide  
; LOCATION: 217..279  
; OTHER INFORMATION: Von Heijne matrix

OTHER INFORMATION: score 5.8172934575094  
; OTHER INFORMATION: seq ALLMAQEVGYLA/GR  
US-09-876-997-47

Query Match 85.6%; Score 15.4; DB 10; Length 1523;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCA 17  
Db 432 CGACTGTGAATCCTCCA 448

RESULT 23  
US-09-935-390A-15  
; Sequence 15, Application US/09935390A  
; Patent No. US20020076761A1  
; GENERAL INFORMATION:  
; APPLICANT: Escobedo, Jaime  
; APPLICANT: Quianjin, Hu  
; APPLICANT: Garcia, Pablo  
; APPLICANT: Williams, Lewis T.  
; TITLE OF INVENTION: Secreted Human Proteins  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/935,390A  
; FILING DATE: 22-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/988,671  
; FILING DATE: 1997-12-11  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane E. R. Potter  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 1369.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 923-2718  
; TELEFAX: (510) 655-3542  
; TLEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1535 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-935-390A-15

Query Match 85.6%; Score 15.4; DB 9; Length 1535;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCA 17  
Db 435 CGACTGTGAATCCTCCA 451

RESULT 24  
US-09-745-763-188  
; Sequence 188, Application US/09745763

Patent No. US20020065394A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
Lavallie, Edward R.  
Collins-Racie, Lisa A.  
Byans, Cheryl  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 188:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1755 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 188:  
US-09-745-763-188  
Query Match 85.6%; Score 15.4; DB 9; Length 1755;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCTCCA 17  
Db 673 CGACTGTGAATCTCCA 689  
RESULT 25  
US-09-925-298-221  
Sequence 221, Application US/09925298  
Publication No. US20020039764A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA103  
CURRENT APPLICATION NUMBER: US/09/925,298  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05881  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 846  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 221  
LENGTH: 1892  
TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1892)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-298-221  
Query Match 85.6%; Score 15.4; DB 9; Length 1892;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCTCCA 17  
Db 763 CGACTGTGAATCTCCA 779  
RESULT 26  
US-10-102-806-221  
Sequence 221, Application US/10102806  
Publication No. US20030054421A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA103P1C1  
CURRENT APPLICATION NUMBER: US/10/102,806  
CURRENT FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: 09/925,298  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05881  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 846  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 221  
LENGTH: 1892  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1892)  
OTHER INFORMATION: n equals a,t,g, or c  
US-10-102-806-221  
Query Match 85.6%; Score 15.4; DB 14; Length 1892;  
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCTCCA 17  
Db 763 CGACTGTGAATCTCCA 779  
RESULT 27  
US-10-322-696-67  
Sequence 67, Application US/10322696  
Publication No. US20040166490A1  
GENERAL INFORMATION:  
APPLICANT: Morris, David W.  
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER  
FILE REFERENCE: 529452001200  
CURRENT APPLICATION NUMBER: US/10/322,696  
CURRENT FILING DATE: 2003-10-17  
NUMBER OF SEQ ID NOS: 186  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 67  
LENGTH: 154504  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)-(154504)

OTHER INFORMATION: n = A,T,C or G  
US-10-322-696-67

Query Match 85.6%; Score 15.4; DB 17; Length 154504;  
Best Local Similarity 94.1%; Pred. No. 4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
|||||  
Db 63891 GACTGTGAATCCTCCAT 63907

RESULT 28  
US-10-672-764A-33  
; Sequence 33, Application US/10672764A  
; Publication No. US20040156832A1  
; GENERAL INFORMATION:  
; APPLICANT: Jolly, Chris  
; TITLE OF INVENTION: Immunoglobulin Compositions and Methods  
; FILE REFERENCE: 1331.1001U  
; CURRENT APPLICATION NUMBER: US/10/672,764A  
; CURRENT FILING DATE: 2003-09-26  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 200000  
; TYPE: DNA  
; ORGANISM: Human  
US-10-672-764A-33

Query Match 85.6%; Score 15.4; DB 17; Length 200000;  
Best Local Similarity 94.1%; Pred. No. 4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
|||||  
Db 107278 GACTGTGAATCCTCCAT 107294

RESULT 29  
US-10-131-786-2  
; Sequence 2, Application US/10131786  
; Publication No. US20030026831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lakkaraju, A.  
; APPLICANT: Dubinsky, J.M.  
; APPLICANT: Low, W.  
; APPLICANT: Rahman, Y.  
; TITLE OF INVENTION: Delivery of antisense oligonucleotides to neurons with anionic li  
; FILE REFERENCE: 600.519US1  
; CURRENT APPLICATION NUMBER: US/10/131,786  
; CURRENT FILING DATE: 2002-04-22  
; PRIOR APPLICATION NUMBER: US 60/285,337  
; PRIOR FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 18  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
US-10-131-786-2

Query Match 83.3%; Score 15; DB 14; Length 18;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CTGTGAATCCTCCAT 18  
|||||  
Db 1 CTGTGAATCCTCCAT 15

RESULT 30  
US-10-202-193-5/c

Sequence 5, Application US/10202193  
; Publication No. US20020192699A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Steinhmann, Kathleen E.  
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
; FILE REFERENCE: PP-01532.103/200130.463D1  
; CURRENT APPLICATION NUMBER: US/10/202,193  
; CURRENT FILING DATE: 2002-07-23  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 1024  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(1024)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-202-193-5

Query Match 83.3%; Score 15; DB 13; Length 1024;  
Best Local Similarity 93.8%; Pred. No. 6.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACTGTGAATCCTCCAT 18  
|||||  
Db 574 ACTGTGAATCCTCCAT 559

RESULT 31  
US-10-417-375-128  
; Sequence 128, Application US/10417375  
; Publication No. US20040219528A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer  
; FILE REFERENCE: 529452001600  
; CURRENT APPLICATION NUMBER: US/10/417,375  
; CURRENT FILING DATE: 2003-04-15  
; NUMBER OF SEQ ID NOS: 176  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 128  
; LENGTH: 430442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-417-375-128

Query Match 83.3%; Score 15; DB 18; Length 430442;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ACTGTGAATCCTCCA 17  
|||||  
Db 139298 ACTGTGAATCCTCCA 139312

RESULT 32  
US-10-425-115-65084/c  
; Sequence 65084, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua



APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 65084  
LENGTH: 225  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_159354C.1  
US-10-425-115-65084

Query Match 82.2%; Score 14.8; DB 18; Length 225;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGATTCCTCCAT 18  
DB 88 CGTCTGTATTCCTCCAT 71

RESULT 33  
US-09-922-293-1999  
Sequence 1999, Application US/09922293  
Publication No. US20040123339A1  
GENERAL INFORMATION:  
APPLICANT: Conner, Timothy W.  
APPLICANT: Heck, Gregory R.  
APPLICANT: Liu, Jingdong  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 16517.254  
CURRENT APPLICATION NUMBER: US/09/922,293  
CURRENT FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/067,000  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: US 60/069,472  
PRIOR FILING DATE: 1997-12-09  
PRIOR APPLICATION NUMBER: US 60/071,479  
PRIOR FILING DATE: 1998-01-13  
PRIOR APPLICATION NUMBER: US 60/074,201  
PRIOR FILING DATE: 1998-02-10  
PRIOR APPLICATION NUMBER: US 60/074,282  
PRIOR FILING DATE: 1998-02-10  
PRIOR APPLICATION NUMBER: US 60/074,280  
PRIOR FILING DATE: 1998-02-10  
PRIOR APPLICATION NUMBER: US 60/074,281  
PRIOR FILING DATE: 1998-02-10  
PRIOR APPLICATION NUMBER: US 60/074,566  
PRIOR FILING DATE: 1998-02-12  
PRIOR APPLICATION NUMBER: US 60/074,567  
PRIOR FILING DATE: 1998-02-12  
PRIOR APPLICATION NUMBER: US 60/074,565  
PRIOR FILING DATE: 1998-02-12  
PRIOR APPLICATION NUMBER: US 60/075,462  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 60/074,789  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 60/075,459  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 60/075,461  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 60/075,464  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 60/075,460  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 60/075,463  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 60/077,231  
PRIOR FILING DATE: 1998-03-09

PRIOR APPLICATION NUMBER: US 60/077,229  
PRIOR FILING DATE: 1998-03-09  
PRIOR APPLICATION NUMBER: US 60/077,230  
PRIOR FILING DATE: 1998-03-09  
PRIOR APPLICATION NUMBER: US 60/078,368  
PRIOR FILING DATE: 1998-03-18  
PRIOR APPLICATION NUMBER: US 60/080,844  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 60/083,067  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: US 60/083,386  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: US 60/083,387  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: US 60/083,388  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: US 60/083,389  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: US 60/085,224  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: US 60/085,223  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: US 60/085,222  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: US 60/085,533  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: US 60/086,186  
PRIOR FILING DATE: 1998-05-21  
PRIOR APPLICATION NUMBER: US 60/086,187  
PRIOR FILING DATE: 1998-05-21  
PRIOR APPLICATION NUMBER: US 60/086,185  
PRIOR FILING DATE: 1998-05-21  
PRIOR APPLICATION NUMBER: US 60/086,184  
PRIOR FILING DATE: 1998-05-21  
PRIOR APPLICATION NUMBER: US 60/086,183  
PRIOR FILING DATE: 1998-05-21  
PRIOR APPLICATION NUMBER: US 60/086,188  
PRIOR FILING DATE: 1998-05-21  
PRIOR APPLICATION NUMBER: US 60/089,524  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: US 60/089,810  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,814  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,808  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,812  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,807  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,806  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,813  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,811  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/089,793  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: US 60/091,405  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/091,247  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/099,667  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: US 60/099,668  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: US 60/099,670  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: US 60/099,697  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: US 60/100,674  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: US 60/100,673

;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: US 60/100,672  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: US 60/100,963  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: US 60/101,131  
;; PRIOR FILING DATE: 1998-09-21  
;; PRIOR APPLICATION NUMBER: US 60/101,132  
;; PRIOR FILING DATE: 1998-09-21  
;; PRIOR APPLICATION NUMBER: US 60/101,130  
;; PRIOR FILING DATE: 1998-09-21  
;; PRIOR APPLICATION NUMBER: US 60/101,508  
;; PRIOR FILING DATE: 1998-09-22  
;; PRIOR APPLICATION NUMBER: US 60/101,344  
;; PRIOR FILING DATE: 1998-09-22  
;; PRIOR APPLICATION NUMBER: US 60/101,347  
;; PRIOR FILING DATE: 1998-09-22  
;; PRIOR APPLICATION NUMBER: US 60/101,343  
;; PRIOR FILING DATE: 1998-09-22  
;; PRIOR APPLICATION NUMBER: US 60/101,707  
;; PRIOR FILING DATE: 1998-09-25  
;; PRIOR APPLICATION NUMBER: US 60/104,126  
;; PRIOR FILING DATE: 1998-10-13  
;; PRIOR APPLICATION NUMBER: US 60/104,128  
;; PRIOR FILING DATE: 1998-10-13  
;; PRIOR APPLICATION NUMBER: US 60/104,127  
;; PRIOR FILING DATE: 1998-10-13  
;; PRIOR APPLICATION NUMBER: US 60/104,124  
;; PRIOR FILING DATE: 1998-10-13  
;; PRIOR APPLICATION NUMBER: US 60/109,018  
;; PRIOR FILING DATE: 1998-11-18  
;; PRIOR APPLICATION NUMBER: US 60/108,996  
;; PRIOR FILING DATE: 1998-11-18  
;; PRIOR APPLICATION NUMBER: US 09/199,129  
;; PRIOR FILING DATE: 1998-11-24  
;; PRIOR APPLICATION NUMBER: US 09/210,297  
;; PRIOR FILING DATE: 1998-12-08  
;; PRIOR APPLICATION NUMBER: US 60/111,981  
;; PRIOR FILING DATE: 1998-12-11  
;; PRIOR APPLICATION NUMBER: US 60/113,224  
;; PRIOR FILING DATE: 1998-12-22  
;; PRIOR APPLICATION NUMBER: US 09/229,413  
;; PRIOR FILING DATE: 1999-01-12  
;; NUMBER OF SEQ ID NOS: 3853  
;; SEQ ID NO 1999  
;; LENGTH: 281  
;; TYPE: DNA  
;; ORGANISM: Glycine max  
US-09-922-293-1999

Query Match 82.2%; Score 14.8; DB 11; Length 281;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCAT 18  
Db 258 CGAGTGTGATCCTCCAT 275  
RESULT 34  
US-10-242-535A-30560/C  
; Sequence 30560, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340

;; PRIOR FILING DATE: 2001-07-13  
;; PRIOR APPLICATION NUMBER: US 60/275,017  
;; PRIOR FILING DATE: 2001-03-12  
;; PRIOR APPLICATION NUMBER: US 60/271,955  
;; PRIOR FILING DATE: 2001-02-28  
;; NUMBER OF SEQ ID NOS: 58994  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 30560  
;; LENGTH: 350  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (340)..(340)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-30560

Query Match 82.2%; Score 14.8; DB 16; Length 350;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCAT 18  
Db 72 CGACTGTGAATCCACAT 55

RESULT 35  
US-10-085-783A-30560/C  
; Sequence 30560, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 30560  
; LENGTH: 350  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (340)..(340)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-085-783A-30560

Query Match 82.2%; Score 14.8; DB 16; Length 350;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCAT 18  
Db 72 CGACTGTGAATCCACAT 55  
RESULT 36  
US-09-954-456-1075  
; Sequence 1075, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance  
; FILE REFERENCE: 689290-76

```
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; SOFTWARE: Patentin version 3.0
; NUMBER OF SEQ ID NOS: 2276
; SEQ ID NO 1075
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1075
```

```
Query Match      82.2%; Score 14.8; DB 9; Length 417;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CGACTGTGAATCCTCCAT 18
Db      278 CGACTGTGAATCCTCCAT 295
```

```
RESULT 37
; Sequence 1774, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Setc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
```

```
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 1774
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1774
```

```
Query Match      82.2%; Score 14.8; DB 9; Length 417;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CGACTGTGAATCCTCCAT 18
Db      278 CGACTGTGAATCCTCCAT 295
```

```
RESULT 38
; Sequence 81, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Logic, Inc.
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 81
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA035540
; NAME/KEY: unsure
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-81
```

```
Query Match      82.2%; Score 14.8; DB 9; Length 417;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CGACTGTGAATCCTCCAT 18
Db      278 CGACTGTGAATCCTCCAT 295
```

```
RESULT 39
; Sequence 19674, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19674
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(474)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19674

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 10; Length 474;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
Db 343 CGACTGTGAATCCTCCAT 326

RESULT 40
US-10-027-632-137650
; Sequence 137650, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137650
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-137650

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 13; Length 578;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
Db 423 CGACTGTGAATCCTCCAT 440

RESULT 41
US-10-027-632-137650
; Sequence 137650, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137650
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-137650

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 15; Length 578;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
Db 423 CGACTGTGAATCCTCCAT 440

RESULT 42
US-10-425-115-47892
; Sequence 47892, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 47892
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_143684C.1
US-10-425-115-47892

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 18; Length 600;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
Db 396 CGACTGTGAATCCTCCAT 413

RESULT 43
US-10-767-701-519/C
; Sequence 519, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovall, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2002-04-30
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
```

FILE REFERENCE: 38-21(5353)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-23  
NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 519  
LENGTH: 645  
TYPE: DNA  
ORGANISM: Sorghum bicolor  
FEATURE:  
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS116219\_1  
US-10-767-701-519

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 17; Length 645;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTCCAT 18  
Db 71 CGATTGTATCTCCAT 54

RESULT 44  
US-10-425-115-108111/c  
Sequence 108111, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(5322)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 108111  
LENGTH: 846  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_30092C.1  
US-10-425-115-108111

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 18; Length 846;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTCCAT 18  
Db 56 CGCTGTGAATCCGCAT 39

RESULT 45  
US-10-260-238-3388  
Sequence 3388, Application US/10260238  
Publication No. US20040016025A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, Paul R.  
APPLICANT: Moughamer, Todd G.  
APPLICANT: Briggs, Steven P.  
APPLICANT: Cooper, Bret  
APPLICANT: Glazebrook, Jane  
APPLICANT: Goff, Stephen A.  
APPLICANT: Katagiri, Fumiyuki  
APPLICANT: Kreps, Joel  
APPLICANT: Provart, Nicholas  
APPLICANT: Rieke, Darrell  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
FILE REFERENCE: 60111-NP  
CURRENT APPLICATION NUMBER: US/10/260,238  
CURRENT FILING DATE: 2002-09-26

PRIOR APPLICATION NUMBER: US 60/325,448  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 3388  
LENGTH: 908  
TYPE: DNA  
ORGANISM: Triticum aestivum  
US-10-260-238-3388

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 16; Length 908;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTCCAT 18  
Db 217 CGAATGTGAATCTCCAT 234

RESULT 46  
US-10-238-075-579/c  
Sequence 579, Application US/10238075  
Publication No. US20030148324A1  
GENERAL INFORMATION:  
APPLICANT: I.N.S.E.R.M.  
TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat  
TITLE OF INVENTION: B.colli, and biological uses of these polynucleotides and of thei  
FILE REFERENCE: BLANDINE  
CURRENT APPLICATION NUMBER: US/10/238,075  
CURRENT FILING DATE: 2002-09-10  
PRIOR APPLICATION NUMBER: 0003145  
PRIOR FILING DATE: 2000-03-10  
NUMBER OF SEQ ID NOS: 1576  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 579  
LENGTH: 966  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-10-238-075-579

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 15; Length 966;  
Best Local Similarity 88.9%; Pred. No. 8.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTCCAT 18  
Db 855 CACTGTGAATCTCCAT 838

RESULT 47  
US-10-437-963-34632  
Sequence 34632, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and uses thereof for Plant Improvement  
FILE REFERENCE: 38-21(5322)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 34632  
LENGTH: 977

TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_38629C.1  
US-10-437-963-34632

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 16; Length 977;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 636 CCACTGTGTATCCTCCAT 653

RESULT 48  
US-10-258-662-58/c

Sequence 58, Application US/10258662  
Publication No. US20040082029A1  
GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: LAL, Preeti

APPLICANT: YUE, Henry

APPLICANT: TANG, Y. Tom

APPLICANT: LU, Dyrung Aina M.

APPLICANT: AZIMZAI, Yalda

APPLICANT: AU-YOUNG, Janice

APPLICANT: HITLMAN, Jennifer L.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: YAO, Monique G.

APPLICANT: BURFORD, Neil

APPLICANT: BATRA, Sateev

APPLICANT: POLICKY, Jennifer J.

TITLE OF INVENTION: RNA METABOLISM PROTEINS

FILE REFERENCE: PR-0771 USN

CURRENT APPLICATION NUMBER: US/10/258, 662

PRIOR FILING DATE: 2002-10-25

PRIOR APPLICATION NUMBER: US 60/201, 875

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/200, 164

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/202, 090

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/210, 232

PRIOR FILING DATE: 2000-06-06

PRIOR APPLICATION NUMBER: US 60/220, 553

PRIOR FILING DATE: 2000-07-25

NUMBER OF SEQ ID NOS: 94

SOFTWARE: PERL Program

SEQ ID NO 58

LENGTH: 1188

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No: 2099829CB1

US-10-258-662-58

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 16; Length 1188;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 666 CGACTGTGAATCCTCCAT 649

RESULT 49  
US-10-424-599-82180/c

Sequence 82180, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424, 599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 82180

LENGTH: 1448

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_45225C.1

US-10-424-599-82180

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 16; Length 1448;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 628 CCACTGTGAATCCTCCAT 611

RESULT 50  
US-10-437-963-35984

Sequence 35984, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437, 963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 35984

LENGTH: 1826

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT4530\_39853C.1

US-10-437-963-35984

Query Match  
Best Local Similarity 82.2%; Score 14.8; DB 16; Length 1826;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 1360 CGACTGTGAATCCTCCAT 1377

Search completed: January 14, 2005, 19:47:52  
Job time : 358.579 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:47:55 ; Search time 63.9474 Seconds  
(without alignments)  
200.074 Million cell updates/sec

Title: US-09-578-453-1  
Sequence: 1 CGACTGTCATCCTCCAT 18

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfilest.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	1415	3	US-09-414-436-2
2	16.4	91.1	1173	2	US-08-184-009-214
3	16.4	91.1	1173	2	US-08-458-356-214
4	16.4	91.1	1173	3	US-08-796-101-45
5	16.4	91.1	1173	3	US-08-460-736-214
6	16.4	91.1	1173	3	US-09-535-370-214
7	16.4	91.1	1173	4	US-09-663-667-214
8	15.4	85.6	396	4	US-08-585-593A-36
9	15.4	85.6	527	4	US-09-621-976-1600
10	15.4	85.6	2516	4	US-09-270-767-1686
11	15	83.3	1024	4	US-09-328-475C-5
12	14.8	82.2	1107	4	US-09-248-796A-8906
13	14.4	80.0	478	4	US-09-621-976-18989
14	14.4	80.0	1093	4	US-09-186-276B-53
15	14.4	80.0	1093	4	US-08-842-445-53
16	14.4	80.0	1093	4	US-09-186-188B-53
17	14.4	80.0	1629	4	US-09-919-039-32
18	14.4	80.0	1850	3	US-08-724-468B-3
19	14.4	80.0	1850	3	US-08-882-164D-3
20	14.4	80.0	3152	4	US-09-399-081A-7
21	14.4	80.0	6911	1	US-08-311-174-4
22	14.4	80.0	8878	1	US-08-205-176-3
23	14	77.8	1395	4	US-09-614-221A-167
24	13.8	76.7	432	4	US-09-513-999C-36262
25	13.8	76.7	514	4	US-09-389-681-417
26	13.8	76.7	514	4	US-09-620-405B-417
27	13.8	76.7	514	4	US-09-433-826B-417

28	13.8	76.7	514	4	US-09-604-287A-417	Sequence 417, App
29	13.8	76.7	514	4	US-09-834-759-417	Sequence 417, App
30	13.8	76.7	514	4	US-09-590-751A-417	Sequence 417, App
31	13.8	76.7	516	4	US-09-489-039A-2781	Sequence 2781, App
32	13.8	76.7	519	4	US-09-389-681-258	Sequence 258, App
33	13.8	76.7	519	4	US-09-620-405B-258	Sequence 258, App
34	13.8	76.7	519	4	US-09-339-338-258	Sequence 258, App
35	13.8	76.7	519	4	US-09-433-826B-258	Sequence 258, App
36	13.8	76.7	519	4	US-09-604-287A-258	Sequence 258, App
37	13.8	76.7	519	4	US-09-834-759-258	Sequence 258, App
38	13.8	76.7	519	4	US-09-590-751A-258	Sequence 258, App
39	13.8	76.7	598	4	US-09-270-767-11839	Sequence 11839, A
40	13.8	76.7	1041	4	US-09-371-307-79	Sequence 79, App
41	13.8	76.7	1146	3	US-09-257-580-1	Sequence 1, App
42	13.8	76.7	1146	3	US-09-257-580-4	Sequence 4, App
43	13.8	76.7	1166	4	US-09-673-395A-57	Sequence 57, App
44	13.8	76.7	1270	4	US-09-270-767-12871	Sequence 12871, A
45	13.8	76.7	1431	4	US-10-000-489-87	Sequence 87, App
46	13.8	76.7	1431	4	US-10-000-489-89	Sequence 89, App
47	13.8	76.7	1458	4	US-09-248-786A-5727	Sequence 5727, App
48	13.8	76.7	1518	4	US-09-719-919A-2	Sequence 2, App
49	13.8	76.7	1734	3	US-08-687-590-60	Sequence 60, App
50	13.8	76.7	1938	4	US-09-673-395A-140	Sequence 140, App
51	13.8	76.7	1955	4	US-09-270-767-13672	Sequence 13672, A
52	13.8	76.7	2070	4	US-09-540-236-1766	Sequence 1766, App
53	13.8	76.7	2307	3	US-09-774-528-314	Sequence 314, App
54	13.8	76.7	4377	3	US-08-911-853-28	Sequence 28, App
55	13.8	76.7	4377	3	US-09-479-409-28	Sequence 28, App
56	13.8	76.7	4377	3	US-09-479-453-28	Sequence 28, App
57	13.8	76.7	5240	3	US-09-171-337A-2	Sequence 2, App
58	13.8	76.7	5240	4	US-09-631-002-24	Sequence 24, App
59	13.8	76.7	89047	4	US-09-596-002-34	Sequence 34, App
60	13.8	76.7	168174	4	US-10-071-411A-63	Sequence 63, App
61	13.8	76.7	168273	4	US-10-071-411A-2	Sequence 2, App
62	13.8	76.7	536165	4	US-09-214-808-1	Sequence 1, App
63	13.4	74.4	79	3	US-08-921-887-42	Sequence 42, App
64	13.4	74.4	392	4	US-09-270-767-27497	Sequence 27497, A
65	13.4	74.4	478	4	US-09-270-767-8209	Sequence 8209, App
66	13.4	74.4	478	4	US-09-270-767-33491	Sequence 23491, A
67	13.4	74.4	600	4	US-09-621-976-10493	Sequence 10493, A
68	13.4	74.4	611	4	US-09-270-767-1675	Sequence 4675, App
69	13.4	74.4	611	4	US-09-270-767-19957	Sequence 19957, A
70	13.4	74.4	769	4	US-09-513-775B-5	Sequence 5, App
71	13.4	74.4	907	4	US-09-513-775B-3	Sequence 3, App
72	13.4	74.4	917	4	US-09-270-767-11849	Sequence 11849, A
73	13.4	74.4	918	1	US-08-328-322-3	Sequence 3, App
74	13.4	74.4	956	4	US-09-641-638-36	Sequence 36, App
75	13.4	74.4	956	4	US-09-641-638-37	Sequence 37, App
76	13.4	74.4	956	4	US-09-641-638-38	Sequence 38, App
77	13.4	74.4	956	4	US-09-641-638-39	Sequence 39, App
78	13.4	74.4	956	4	US-09-641-638-40	Sequence 40, App
79	13.4	74.4	956	4	US-10-170-097-36	Sequence 36, App
80	13.4	74.4	956	4	US-10-170-097-37	Sequence 37, App
81	13.4	74.4	956	4	US-10-170-097-38	Sequence 38, App
82	13.4	74.4	956	4	US-10-170-097-39	Sequence 39, App
83	13.4	74.4	956	4	US-10-170-097-40	Sequence 40, App
84	13.4	74.4	1001	4	US-09-641-638-428	Sequence 628, App
85	13.4	74.4	1101	4	US-10-170-097-428	Sequence 628, App
86	13.4	74.4	1152	4	US-09-107-532A-2119	Sequence 2119, App
87	13.4	74.4	1158	4	US-09-328-352-2024	Sequence 2024, App
88	13.4	74.4	1191	3	US-09-134-001C-1709	Sequence 1709, App
89	13.4	74.4	1206	4	US-09-248-786A-3	Sequence 3, App
90	13.4	74.4	1222	1	US-08-358-352-1	Sequence 1, App
91	13.4	74.4	1281	4	US-09-620-312D-849	Sequence 849, App
92	13.4	74.4	1418	4	US-09-673-395A-26	Sequence 26, App
93	13.4	74.4	1430	2	US-08-743-130A-1	Sequence 1, App
94	13.4	74.4	2024	4	US-09-398-395A-25	Sequence 25, App
95	13.4	74.4	2024	4	US-09-887-586A-25	Sequence 25, App
96	13.4	74.4	2024	4	US-09-895-752-25	Sequence 25, App
97	13.4	74.4	2024	4	US-09-903-912B-25	Sequence 25, App
98	13.4	74.4	2024	4	US-09-900-797-25	Sequence 25, App
99	13.4	74.4	2025	2	US-08-937-540-1	Sequence 1, App
100	13.4	74.4	2745	1	US-08-328-322-6	Sequence 6, App

101	13.4	74.4	2748	4	US-09-614-221A-288	Sequence 288, App	174	12.8	71.1	36	1	US-08-403-762A-153	Sequence 153, App
102	13.4	74.4	3288	1	US-08-208-008C-3	Sequence 3, Appl1	175	12.8	71.1	61	1	US-07-744-282C-111	Sequence 111, App
103	13.4	74.4	3500	2	US-08-328-322-4	Sequence 26, Appl1	176	12.8	71.1	61	1	US-07-744-282C-113	Sequence 113, App
104	13.4	74.4	11444	2	US-08-222-617A-26	Sequence 1, Appl1	177	12.8	71.1	61	5	PCT-US92-06821A-57	Sequence 57, Appl1
105	13.4	74.4	12364	2	US-08-222-617A-1	Sequence 1, Appl1	178	12.8	71.1	61	5	PCT-US92-06821A-59	Sequence 59, Appl1
106	13.4	74.4	20674	4	US-09-641-638-651	Sequence 651, App	179	12.8	71.1	236	4	US-09-513-999C-12376	Sequence 12376, A
107	13.4	74.4	20674	4	US-10-170-097-651	Sequence 651, App	180	12.8	71.1	312	3	US-09-188-930-21	Sequence 21, Appl1
108	13.4	74.4	41708	3	US-09-470-512A-3	Sequence 3, Appl1	181	12.8	71.1	312	4	US-09-312-283C-21	Sequence 21, Appl1
109	13.4	74.4	41708	4	US-09-676-539-18	Sequence 18, Appl1	182	12.8	71.1	315	3	US-09-060-756-127	Sequence 127, App
110	13.4	74.4	99916	4	US-09-816-095-3	Sequence 3, Appl1	183	12.8	71.1	324	4	US-09-670-314-127	Sequence 127, App
111	13.4	74.4	786431	4	US-09-751-389-3	Sequence 3, Appl1	184	12.8	71.1	324	4	US-09-328-352-995	Sequence 995, App
112	13.2	73.3	246	4	US-09-107-532A-525	Sequence 971, App	185	12.8	71.1	335	3	US-08-905-223-194	Sequence 194, App
113	13.2	73.3	290	4	US-09-313-294A-971	Sequence 5, Appl1	186	12.8	71.1	351	4	US-09-513-999C-30457	Sequence 30457, A
114	13.2	73.3	474	1	US-07-845-592J-5	Sequence 213, App	187	12.8	71.1	360	4	US-09-543-681A-6	Sequence 6, Appl1
115	13.2	73.3	657	4	US-09-084-303B-213	Sequence 3102, App	188	12.8	71.1	370	4	US-09-513-999C-24920	Sequence 24920, A
116	13.2	73.3	780	4	US-09-246-796A-3102	Sequence 25, Appl1	189	12.8	71.1	394	4	US-09-513-999C-27193	Sequence 27193, A
117	13.2	73.3	831	4	US-09-904-615-25	Sequence 657, App	190	12.8	71.1	398	4	US-09-621-976-10095	Sequence 10095, A
118	13.2	73.3	867	4	US-09-252-991A-6957	Sequence 2240, App	191	12.8	71.1	409	4	US-09-270-767-2400	Sequence 2400, App
119	13.2	73.3	888	4	US-09-134-000C-2240	Sequence 2, Appl1	192	12.8	71.1	409	4	US-09-270-767-17682	Sequence 17682, A
120	13.2	73.3	1166	1	US-07-845-592J-3	Sequence 3, Appl1	193	12.8	71.1	423	4	US-09-248-796A-2788	Sequence 2788, App
121	13.2	73.3	1443	4	US-09-802-927-2	Sequence 2, Appl1	194	12.8	71.1	424	4	US-09-621-976-17245	Sequence 17245, A
122	13.2	73.3	1444	4	US-09-612-402B-11	Sequence 11, Appl1	195	12.8	71.1	446	4	US-09-621-976-18000	Sequence 18000, A
123	13.2	73.3	1614	4	US-09-134-000C-926	Sequence 926, App	196	12.8	71.1	510	4	US-09-621-976-17134	Sequence 17134, A
124	13.2	73.3	1656	4	US-09-248-796A-608	Sequence 608, App	197	12.8	71.1	510	4	US-09-621-976-17134	Sequence 17134, A
125	13.2	73.3	1734	4	US-09-248-796A-4005	Sequence 4005, App	198	12.8	71.1	513	4	US-09-248-796A-6137	Sequence 6137, App
126	13.2	73.3	2038	3	US-09-323-872A-45	Sequence 4, Appl1	199	12.8	71.1	525	4	US-09-270-767-5250	Sequence 5250, App
127	13.2	73.3	2052	4	US-08-869-423-6	Sequence 6, Appl1	200	12.8	71.1	525	4	US-09-270-767-20532	Sequence 20532, A
128	13.2	73.3	2116	4	US-08-635-552A-1	Sequence 1, Appl1	201	12.8	71.1	622	4	US-09-621-976-17866	Sequence 17866, App
129	13.2	73.3	2117	3	US-09-155-768-1	Sequence 1, Appl1	202	12.8	71.1	632	4	US-09-270-767-3402	Sequence 3402, App
130	13.2	73.3	2235	4	US-09-248-796A-3987	Sequence 3987, App	203	12.8	71.1	632	4	US-09-270-767-18684	Sequence 18684, A
131	13.2	73.3	2520	4	US-09-799-451-4	Sequence 4, Appl1	204	12.8	71.1	655	3	US-09-248-796A-2789	Sequence 2789, App
132	13.2	73.3	2601	4	US-09-799-451-3	Sequence 3, Appl1	205	12.8	71.1	655	3	US-09-328-111-727	Sequence 727, App
133	13.2	73.3	2728	4	US-09-620-312D-572	Sequence 572, App	206	12.8	71.1	667	4	US-09-328-111-7197	Sequence 7197, App
134	13.2	73.3	2949	4	US-09-556-877-170	Sequence 170, App	207	12.8	71.1	667	4	US-09-270-767-7197	Sequence 22479, A
135	13.2	73.3	2949	4	US-09-620-412C-170	Sequence 170, App	208	12.8	71.1	711	6	5169941-24	Patent No. 5169941
136	13.2	73.3	2949	4	US-09-598-419-170	Sequence 170, App	209	12.8	71.1	711	6	US-08-470-535-13	Sequence 13, Appl1
137	13.2	73.3	3021	4	US-09-556-877-182	Sequence 182, App	210	12.8	71.1	729	3	US-09-270-767-77857	Sequence 77857, App
138	13.2	73.3	3021	4	US-09-620-412C-182	Sequence 182, App	211	12.8	71.1	729	3	US-09-270-767-77857	Sequence 77857, App
139	13.2	73.3	3021	4	US-09-598-419-182	Sequence 21, Appl1	212	12.8	71.1	736	4	US-09-270-767-23039	Sequence 23039, A
140	13.2	73.3	3324	4	US-09-612-402B-24	Sequence 24, Appl1	213	12.8	71.1	737	4	US-09-270-767-14611	Sequence 14611, A
141	13.2	73.3	3354	4	US-09-612-402B-23	Sequence 23, Appl1	214	12.8	71.1	737	4	US-09-673-395A-55	Sequence 55, Appl1
142	13.2	73.3	3529	4	US-09-799-451-721	Sequence 7, Appl1	215	12.8	71.1	774	4	US-09-107-532A-2311	Sequence 2311, App
143	13.2	73.3	3738	1	US-08-158-232-7	Sequence 7, Appl1	216	12.8	71.1	786	2	US-08-644-6649-29	Sequence 29, Appl1
144	13.2	73.3	3738	1	US-08-304-626-7	Sequence 7, Appl1	217	12.8	71.1	786	2	US-08-761-277A-29	Sequence 29, Appl1
145	13.2	73.3	3738	3	US-08-611-928-7	Sequence 7, Appl1	218	12.8	71.1	789	4	US-09-248-796A-2661	Sequence 2661, App
146	13.2	73.3	3738	3	US-09-173-891-7	Sequence 1, Appl1	219	12.8	71.1	798	6	5169941-23	Patent No. 5169941
147	13.2	73.3	4354	3	US-09-802-927-1	Sequence 1, Appl1	220	12.8	71.1	813	4	US-09-107-532A-3565	Sequence 3565, App
148	13.2	73.3	4415	3	US-09-486-580A-1	Sequence 1, Appl1	221	12.8	71.1	817	4	US-09-799-451-197	Sequence 197, App
149	13.2	73.3	4435	4	US-09-612-402B-1	Sequence 1, Appl1	222	12.8	71.1	823	3	US-08-998-416-505	Sequence 505, App
150	13.2	73.3	7431	3	US-09-306-998-2	Sequence 2, Appl1	223	12.8	71.1	861	4	US-09-756-983-23	Sequence 23, Appl1
151	13.2	73.3	7650	4	US-09-221-017B-911	Sequence 911, App	224	12.8	71.1	863	4	US-09-673-395A-540	Sequence 540, App
152	13.2	73.3	45716	3	US-08-965-048-5	Sequence 5, Appl1	225	12.8	71.1	870	4	US-09-107-532A-141	Sequence 141, App
153	13.2	73.3	45989	3	US-08-965-048-6	Sequence 6, Appl1	226	12.8	71.1	913	2	US-08-975-316-61	Sequence 61, Appl1
154	13.2	73.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl1	227	12.8	71.1	913	4	US-09-615-192A-61	Sequence 61, Appl1
155	13.2	73.3	4403765	3	US-09-103-840A-1	Sequence 1, Appl1	228	12.8	71.1	913	4	US-09-169-789-61	Sequence 61, Appl1
156	13.2	73.3	4411529	3	US-09-103-840A-2	Sequence 2, Appl1	229	12.8	71.1	997	1	US-08-480-784-25	Sequence 25, Appl1
157	13.2	73.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl1	230	12.8	71.1	997	1	US-08-483-553-25	Sequence 25, Appl1
158	13	72.2	544	4	US-09-673-395A-118	Sequence 118, App	231	12.8	71.1	997	1	US-08-487-002-25	Sequence 25, Appl1
159	13	72.2	749	4	US-09-673-395A-48	Sequence 48, Appl1	232	12.8	71.1	997	1	US-08-483-554B-25	Sequence 25, Appl1
160	13	72.2	889	4	US-10-140-002-41	Sequence 41, Appl1	233	12.8	71.1	997	1	US-08-488-011B-25	Sequence 25, Appl1
161	13	72.2	1027	4	US-09-684-405-8	Sequence 8, Appl1	234	12.8	71.1	997	3	US-08-850-727-25	Sequence 25, Appl1
162	13	72.2	1825	4	US-09-270-767-14929	Sequence 14929, App	235	12.8	71.1	997	5	PCT-US95-10203-25	Sequence 25, Appl1
163	13	72.2	3001	4	US-09-539-333D-180	Sequence 180, App	236	12.8	71.1	997	5	PCT-US95-10203-25	Sequence 25, Appl1
164	12.8	71.1	19	3	US-08-911-853-5	Sequence 5, Appl1	237	12.8	71.1	997	5	PCT-US95-10203-25	Sequence 25, Appl1
165	12.8	71.1	19	3	US-09-479-409-5	Sequence 5, Appl1	238	12.8	71.1	1089	4	US-09-248-796A-7664	Sequence 7664, App
166	12.8	71.1	19	3	US-09-479-409-5	Sequence 5, Appl1	239	12.8	71.1	1101	4	US-09-248-796A-4873	Sequence 4873, App
167	12.8	71.1	21	4	US-08-649-950-56	Sequence 56, Appl1	240	12.8	71.1	1116	1	US-08-244-646-14	Sequence 14, Appl1
168	12.8	71.1	23	3	US-09-242-632A-3	Sequence 3, Appl1	241	12.8	71.1	1141	3	US-09-147-550-10	Sequence 10, Appl1
169	12.8	71.1	34	5	PCT-US95-07554-1	Sequence 1, Appl1	242	12.8	71.1	1141	3	US-09-557-917-10	Sequence 10, Appl1
170	12.8	71.1	35	1	US-07-744-282C-110	Sequence 110, App	243	12.8	71.1	1144	4	US-09-257-179-15	Sequence 15, Appl1
171	12.8	71.1	35	1	US-07-744-282C-112	Sequence 112, App	244	12.8	71.1	1160	6	5169941-2	Patent No. 5169941
172	12.8	71.1	35	1	PCT-US92-06821A-56	Sequence 56, Appl1	245	12.8	71.1	1192	4	US-09-566-921-122	Sequence 122, App
173	12.8	71.1	35	5	PCT-US92-06821A-58	Sequence 58, Appl1	246	12.8	71.1	1200	4	US-09-614-221A-31	Sequence 31, Appl1



C 247	12.8	71.1	1233	4	US-09-543-681A-99	Sequence 99, Appl	320	12.8	71.1	4594	3	US-08-796-101-9	Sequence 9, Appl
C 248	12.8	71.1	1249	4	US-09-495-050A-187	Sequence 187, App	321	12.8	71.1	4594	3	US-09-085-273-43	Sequence 43, Appl
C 249	12.8	71.1	1283	4	US-09-270-767-11737	Sequence 11737, A	322	12.8	71.1	4594	3	US-09-085-273-45	Sequence 45, Appl
C 250	12.8	71.1	1307	2	US-08-078-311-17	Sequence 17, Appl	323	12.8	71.1	4594	4	US-09-916-963-43	Sequence 43, Appl
C 251	12.8	71.1	1307	2	US-08-460-402-17	Sequence 17, Appl	324	12.8	71.1	4594	4	US-09-916-963-45	Sequence 45, Appl
C 252	12.8	71.1	1344	4	US-09-328-352-3791	Sequence 3791, Ap	325	12.8	71.1	4768	2	US-08-658-665-41	Sequence 41, Appl
C 253	12.8	71.1	1350	4	US-09-799-451-799	Sequence 799, App	326	12.8	71.1	4768	3	US-08-796-101-5	Sequence 5, Appl
C 254	12.8	71.1	1365	4	US-09-023-655-119	Sequence 119, App	327	12.8	71.1	4768	3	US-09-085-273-41	Sequence 41, Appl
C 255	12.8	71.1	1412	4	US-09-615-192A-129	Sequence 129, App	328	12.8	71.1	4768	3	US-09-916-963-41	Sequence 41, Appl
C 256	12.8	71.1	1412	4	US-09-169-789-129	Sequence 129, App	329	12.8	71.1	5062	2	US-08-658-665-187	Sequence 187, App
C 257	12.8	71.1	1448	4	US-09-599-360B-3	Sequence 3, Appl	330	12.8	71.1	5183	2	US-08-796-101-42	Sequence 42, Appl
C 258	12.8	71.1	1448	4	US-09-599-360B-49	Sequence 49, Appl	331	12.8	71.1	5183	2	US-08-459-568-3	Sequence 3, Appl
C 259	12.8	71.1	1523	1	US-08-184-327A-7	Sequence 7, Appl	332	12.8	71.1	5183	2	US-08-399-411-3	Sequence 3, Appl
C 260	12.8	71.1	1523	5	PCT-US95-00670-7	Sequence 7, Appl	333	12.8	71.1	5275	3	US-08-796-101-49	Sequence 49, Appl
C 261	12.8	71.1	1619	4	US-09-620-312D-749	Sequence 749, App	334	12.8	71.1	5359	2	US-09-070-060-1	Sequence 1, Appl
C 262	12.8	71.1	1624	4	US-10-140-002-181	Sequence 181, App	335	12.8	71.1	5372	2	US-09-357-746-1	Sequence 1, Appl
C 263	12.8	71.1	1665	4	US-09-328-352-3670	Sequence 3670, Ap	336	12.8	71.1	5512	4	US-09-866-153-5	Sequence 5, Appl
C 264	12.8	71.1	1707	4	US-09-484-970B-13	Sequence 13, Appl	337	12.8	71.1	5512	4	US-09-693-467A-5	Sequence 5, Appl
C 265	12.8	71.1	1800	4	US-09-248-796A-9868	Sequence 9868, Ap	338	12.8	71.1	5512	4	US-09-270-976-5	Sequence 5, Appl
C 266	12.8	71.1	1830	4	US-09-149-476-129	Sequence 129, App	339	12.8	71.1	5562	4	US-10-204-708-63	Sequence 63, Appl
C 267	12.8	71.1	1879	4	US-09-023-655-742	Sequence 742, App	340	12.8	71.1	5741	1	US-07-706-639-4	Sequence 4, Appl
C 268	12.8	71.1	1929	4	US-09-040-229B-7	Sequence 7, Appl	341	12.8	71.1	5741	1	US-07-998-931-4	Sequence 4, Appl
C 269	12.8	71.1	1971	4	US-09-248-796A-6134	Sequence 6134, Ap	342	12.8	71.1	5868	3	US-08-516-859A-3	Sequence 3, Appl
C 270	12.8	71.1	2025	4	US-09-248-796A-11264	Sequence 11264, A	343	12.8	71.1	5868	3	US-09-586-472-3	Sequence 3, Appl
C 271	12.8	71.1	2045	1	US-08-184-327A-1	Sequence 1, Appl	344	12.8	71.1	5868	4	US-09-528-706-3	Sequence 3, Appl
C 272	12.8	71.1	2046	5	PCT-US95-00670-1	Sequence 1, Appl	345	12.8	71.1	6660	2	US-08-771-602D-1	Sequence 1, Appl
C 273	12.8	71.1	2055	4	US-09-023-655-1036	Sequence 1036, Ap	346	12.8	71.1	6660	3	US-09-232-446B-1	Sequence 1, Appl
C 274	12.8	71.1	2062	1	US-08-073-383-5	Sequence 5, Appl	347	12.8	71.1	7326	1	US-08-194-468-1	Sequence 1, Appl
C 275	12.8	71.1	2062	3	US-08-328-239A-3	Sequence 3, Appl	348	12.8	71.1	7326	3	US-09-514-247A-7	Sequence 7, Appl
C 276	12.8	71.1	2062	5	PCT-US94-06365-5	Sequence 5, Appl	349	12.8	71.1	7326	4	US-09-686-316-1	Sequence 1, Appl
C 277	12.8	71.1	2062	5	PCT-US95-13661-3	Sequence 3, Appl	350	12.8	71.1	7344	3	US-08-961-739-1	Sequence 1, Appl
C 278	12.8	71.1	2115	4	US-09-614-221A-111	Sequence 11, App	351	12.8	71.1	11014	4	US-08-956-711E-91	Sequence 91, Appl
C 279	12.8	71.1	2133	4	US-08-427-993B-2	Sequence 2, Appl	352	12.8	71.1	11014	4	US-08-781-986A-91	Sequence 91, Appl
C 280	12.8	71.1	2333	2	US-08-478-609A-2	Sequence 2, Appl	353	12.8	71.1	11464	3	US-07-906-870A-2	Sequence 2, Appl
C 281	12.8	71.1	2366	4	US-09-270-767-13819	Sequence 13819, A	354	12.8	71.1	11327	1	US-07-906-870A-2	Sequence 2, Appl
C 282	12.8	71.1	2429	6	5248599-1	Patent No. 5248599	355	12.8	71.1	126019	4	US-09-596-002-14	Sequence 14, Appl
C 283	12.8	71.1	2429	4	US-09-023-655-76	Sequence 76, Appl	356	12.8	71.1	126019	4	US-09-819-607-3	Sequence 3, Appl
C 284	12.8	71.1	2504	4	US-09-620-312D-778	Sequence 778, App	357	12.8	71.1	13675	4	US-08-311-721A-135	Sequence 135, App
C 285	12.8	71.1	2504	4	US-09-620-312D-778	Sequence 778, App	358	12.8	71.1	14453	3	US-09-146-033-5	Sequence 5, Appl
C 286	12.8	71.1	2652	4	US-09-489-039A-4078	Sequence 4078, Ap	359	12.8	71.1	55827	4	US-09-813-133A-3	Sequence 3, Appl
C 287	12.8	71.1	2753	2	US-07-952-853-5	Sequence 5, Appl	360	12.8	71.1	1230025	4	US-09-198-452A-1	Sequence 1, Appl
C 288	12.8	71.1	2753	2	US-08-914-848-5	Sequence 5, Appl	361	12.8	71.1	1664976	4	US-08-916-421B-1	Sequence 1, Appl
C 289	12.8	71.1	2884	4	US-10-101-464A-291	Sequence 291, App	362	12.8	71.1	1664976	4	US-09-692-570-1	Sequence 1, Appl
C 290	12.8	71.1	2889	4	US-08-537-002A-4	Sequence 4, Appl	363	12.4	68.9	19	4	US-09-696-791-3731	Sequence 3731, App
C 291	12.8	71.1	2889	3	US-08-863-010-4	Sequence 4, Appl	364	12.4	68.9	21	6	5256770-34	Patent No. 5256770
C 292	12.8	71.1	2889	3	US-09-024-429-4	Sequence 4, Appl	365	12.4	68.9	21	6	5466668-37	Patent No. 5466668
C 293	12.8	71.1	2901	4	US-09-342-325C-43	Sequence 43, Appl	366	12.4	68.9	29	3	US-08-838-151A-42	Sequence 42, Appl
C 294	12.8	71.1	2917	1	US-08-592-936B-20	Sequence 20, Appl	367	12.4	68.9	39	3	US-08-838-151A-22	Sequence 22, Appl
C 295	12.8	71.1	2917	2	US-09-111-573-20	Sequence 20, Appl	368	12.4	68.9	51	4	US-09-443-199C-226	Sequence 226, App
C 296	12.8	71.1	3054	4	US-09-799-451-247	Sequence 247, App	369	12.4	68.9	59	3	US-08-965-762-28	Sequence 28, Appl
C 297	12.8	71.1	3081	4	US-09-248-796A-2025	Sequence 2025, Ap	370	12.4	68.9	59	4	US-09-911-927-28	Sequence 28, Appl
C 298	12.8	71.1	3149	4	US-09-392-714-3	Sequence 3, Appl	371	12.4	68.9	59	4	US-09-911-882-28	Sequence 28, Appl
C 299	12.8	71.1	3225	4	US-08-658-665-54	Sequence 54, Appl	372	12.4	68.9	87	3	US-08-976-413A-325	Sequence 325, App
C 300	12.8	71.1	3409	2	US-08-658-665-54	Sequence 54, Appl	373	12.4	68.9	168	4	US-09-513-999C-23298	Sequence 23298, A
C 301	12.8	71.1	3409	3	US-08-796-101-18	Sequence 18, Appl	374	12.4	68.9	168	4	US-09-513-999C-10476	Sequence 10476, A
C 302	12.8	71.1	3409	3	US-09-085-273-54	Sequence 54, Appl	375	12.4	68.9	242	4	US-09-513-999C-35084	Sequence 35084, A
C 303	12.8	71.1	3409	4	US-09-916-963-54	Sequence 54, Appl	376	12.4	68.9	266	4	US-09-313-294A-1595	Sequence 1595, App
C 304	12.8	71.1	3499	2	US-08-658-665-52	Sequence 52, Appl	377	12.4	68.9	267	4	US-09-313-294A-1595	Sequence 1595, App
C 305	12.8	71.1	3499	3	US-08-796-101-16	Sequence 16, Appl	378	12.4	68.9	312	4	US-09-710-279-1093	Sequence 1093, App
C 306	12.8	71.1	3499	4	US-09-085-273-52	Sequence 52, Appl	379	12.4	68.9	340	4	US-09-270-767-8658	Sequence 8658, App
C 307	12.8	71.1	3499	4	US-09-916-963-52	Sequence 52, Appl	380	12.4	68.9	340	4	US-09-270-767-8658	Sequence 8658, App
C 308	12.8	71.1	3600	1	US-08-537-002A-5	Sequence 5, Appl	381	12.4	68.9	340	4	US-09-270-767-23940	Sequence 23940, A
C 309	12.8	71.1	3600	3	US-08-863-010-5	Sequence 5, Appl	382	12.4	68.9	398	4	US-09-513-999C-16838	Sequence 16838, A
C 310	12.8	71.1	3600	3	US-09-024-429-5	Sequence 5, Appl	383	12.4	68.9	400	4	US-08-956-111E-2930	Sequence 2930, App
C 311	12.8	71.1	3600	4	US-10-101-464A-158	Sequence 458, App	384	12.4	68.9	400	4	US-08-781-986A-2930	Sequence 2930, App
C 312	12.8	71.1	3684	4	US-09-205-258-115	Sequence 115, App	385	12.4	68.9	408	3	US-09-134-001C-2068	Sequence 2068, App
C 313	12.8	71.1	4338	1	US-08-015-986A-1	Sequence 1, Appl	386	12.4	68.9	450	4	US-09-621-976-1756	Sequence 1756, App
C 314	12.8	71.1	4338	2	US-08-446-363-1	Sequence 1, Appl	387	12.4	68.9	532	3	US-08-998-416-145	Sequence 145, App
C 315	12.8	71.1	4396	3	US-09-821-736-1	Sequence 1, Appl	388	12.4	68.9	551	4	US-09-311-021-153	Sequence 153, App
C 316	12.8	71.1	4468	4	US-09-620-312D-243	Sequence 243, App	389	12.4	68.9	564	4	US-09-710-279-2225	Sequence 2225, App
C 317	12.8	71.1	4594	2	US-08-658-665-43	Sequence 43, Appl	390	12.4	68.9	594	3	US-08-998-416-143	Sequence 143, App
C 318	12.8	71.1	4594	2	US-08-658-665-45	Sequence 45, Appl	391	12.4	68.9	598	3	US-09-328-111-241	Sequence 241, App
C 319	12.8	71.1	4594	3	US-08-796-101-7	Sequence 7, Appl	392	12.4	68.9	620	4	US-09-270-767-4356	Sequence 4356, App

C 333	12.4	68.9	620	4	US-09-270-767-19638	Sequence 19638, A
C 334	12.4	68.9	623	3	US-09-286-529-9	Sequence 9, Appli
C 335	12.4	68.9	627	4	US-09-710-279-1403	Sequence 1403, Ap
C 336	12.4	68.9	642	3	US-09-134-001C-1872	Sequence 1872, Ap
C 337	12.4	68.9	683	4	US-10-101-464A-315	Sequence 315, App
C 338	12.4	68.9	700	4	US-09-702-705-184	Sequence 184, App
C 339	12.4	68.9	700	4	US-09-736-457-184	Sequence 184, App
C 400	12.4	68.9	700	4	US-09-614-124B-184	Sequence 184, App
C 401	12.4	68.9	700	4	US-09-671-325-184	Sequence 184, App
C 402	12.4	68.9	700	4	US-09-589-184-184	Sequence 184, App
C 403	12.4	68.9	700	4	US-09-658-824-184	Sequence 184, App
C 404	12.4	68.9	724	3	US-08-998-416-627	Sequence 627, App
C 405	12.4	68.9	831	2	US-08-743-637B-188	Sequence 188, App
C 406	12.4	68.9	885	4	US-09-248-796A-6594	Sequence 6594, Ap
C 407	12.4	68.9	893	3	US-09-286-529-8	Sequence 8, Appli
C 408	12.4	68.9	1011	3	US-09-095-163-1	Sequence 1, Appli
C 409	12.4	68.9	1029	3	US-09-485-648-5	Sequence 5, Appli
C 410	12.4	68.9	1029	3	US-09-503-565-5	Sequence 5, Appli
C 411	12.4	68.9	1029	4	US-09-485-649-5	Sequence 5, Appli
C 412	12.4	68.9	1049	4	US-09-270-767-14935	Sequence 14935, A
C 413	12.4	68.9	1073	3	US-08-960-780-37	Sequence 37, Appli
C 414	12.4	68.9	1073	3	US-09-073-898-37	Sequence 37, Appli
C 415	12.4	68.9	1073	4	US-09-850-351A-37	Sequence 37, Appli
C 416	12.4	68.9	1104	4	US-09-328-352-877	Sequence 877, App
C 417	12.4	68.9	1128	4	US-09-134-000C-1576	Sequence 1576, Ap
C 418	12.4	68.9	1132	4	US-08-956-171E-868	Sequence 868, App
C 419	12.4	68.9	1132	4	US-08-781-986A-868	Sequence 868, App
C 420	12.4	68.9	1170	3	US-08-467-023-94	Sequence 94, Appli
C 421	12.4	68.9	1278	3	US-08-467-023-96	Sequence 96, Appli
C 422	12.4	68.9	1284	4	US-09-246-796A-6479	Sequence 6479, Ap
C 423	12.4	68.9	1332	4	US-09-270-767-11937	Sequence 11937, A
C 424	12.4	68.9	1337	3	US-08-467-023-1	Sequence 1, Appli
C 425	12.4	68.9	1456	4	US-09-056-285A-6	Sequence 6, Appli
C 426	12.4	68.9	1460	4	US-09-774-528-413	Sequence 413, App
C 427	12.4	68.9	1473	4	US-09-056-285A-9	Sequence 9, Appli
C 428	12.4	68.9	1476	4	US-09-710-279-1797	Sequence 1797, Ap
C 429	12.4	68.9	1514	3	US-09-162-524-4	Sequence 4, Appli
C 430	12.4	68.9	1542	4	US-09-543-681A-654	Sequence 654, App
C 431	12.4	68.9	1566	4	US-09-489-039A-879	Sequence 879, App
C 432	12.4	68.9	1612	4	US-09-270-767-14918	Sequence 14918, A
C 433	12.4	68.9	1620	3	US-09-134-001C-2281	Sequence 2281, Ap
C 434	12.4	68.9	1635	4	US-09-248-796A-12936	Sequence 12936, A
C 435	12.4	68.9	1652	4	US-09-976-594-809	Sequence 809, App
C 436	12.4	68.9	1658	4	US-09-016-434-1320	Sequence 1320, Ap
C 437	12.4	68.9	1659	4	US-09-248-796A-6048	Sequence 6048, Ap
C 438	12.4	68.9	1664	4	US-09-397-787-68	Sequence 68, Appli
C 439	12.4	68.9	1664	4	US-09-397-787-267	Sequence 267, App
C 440	12.4	68.9	1795	4	US-09-799-451-342	Sequence 342, App
C 441	12.4	68.9	1827	4	US-09-583-110-1663	Sequence 1663, App
C 442	12.4	68.9	1882	3	US-08-501-572-4	Sequence 4, Appli
C 443	12.4	68.9	1882	3	US-09-040-444-4	Sequence 4, Appli
C 444	12.4	68.9	1931	4	US-09-429-906B-5	Sequence 5, Appli
C 445	12.4	68.9	1946	4	US-09-566-921-123	Sequence 123, App
C 446	12.4	68.9	1947	3	US-09-134-001C-2324	Sequence 2324, Ap
C 447	12.4	68.9	2201	3	US-09-330-970-2	Sequence 2, Appli
C 448	12.4	68.9	2311	4	US-09-774-528-35	Sequence 35, Appli
C 449	12.4	68.9	2403	3	US-09-134-001C-2818	Sequence 2818, Ap
C 450	12.4	68.9	2427	4	US-09-489-039A-4376	Sequence 4376, Ap
C 451	12.4	68.9	2448	4	US-09-799-451-929	Sequence 929, App
C 452	12.4	68.9	2447	4	US-08-387-707-12	Sequence 12, Appli
C 453	12.4	68.9	2447	4	US-08-405-271A-12	Sequence 12, Appli
C 454	12.4	68.9	2478	3	US-09-215-694-31	Sequence 31, Appli
C 455	12.4	68.9	2707	4	US-09-870-956-1	Sequence 1, Appli
C 456	12.4	68.9	2783	4	US-09-870-956-55	Sequence 55, Appli
C 457	12.4	68.9	2783	4	US-09-620-312D-107	Sequence 107, App
C 458	12.4	68.9	2799	4	US-08-887-534A-44	Sequence 44, Appli
C 459	12.4	68.9	2799	4	US-09-527-431-44	Sequence 44, Appli
C 460	12.4	68.9	2799	4	US-09-446-861-44	Sequence 44, Appli
C 461	12.4	68.9	2826	3	US-09-010-998-1	Sequence 1, Appli
C 462	12.4	68.9	2843	3	US-09-620-312D-211	Sequence 211, App
C 463	12.4	68.9	3102	4	US-09-710-279-3682	Sequence 3682, Ap
C 464	12.4	68.9	3150	3	US-09-522-666-5	Sequence 5, Appli
C 465	12.4	68.9	3177	4	US-09-710-279-3717	Sequence 3717, Ap

C 466	12.4	68.9	3247	3	US-08-718-388-4	Sequence 4, Appli
C 467	12.4	68.9	3320	4	US-09-710-279-3649	Sequence 3649, Ap
C 468	12.4	68.9	3390	4	US-09-252-991A-13356	Sequence 13356, A
C 469	12.4	68.9	3663	3	US-09-499-884-11	Sequence 11, Appli
C 470	12.4	68.9	3767	4	US-09-710-279-4213	Sequence 4213, Ap
C 471	12.4	68.9	3830	4	US-09-710-279-3773	Sequence 3773, Ap
C 472	12.4	68.9	4079	4	US-09-016-434-1412	Sequence 1412, Ap
C 473	12.4	68.9	4487	4	US-09-484-970B-96	Sequence 96, Appli
C 474	12.4	68.9	4858	4	US-08-956-171E-379	Sequence 379, App
C 475	12.4	68.9	4858	4	US-08-781-986A-379	Sequence 379, App
C 476	12.4	68.9	5053	4	US-09-620-312D-590	Sequence 590, App
C 477	12.4	68.9	5197	2	US-08-131-365B-53	Sequence 53, Appli
C 478	12.4	68.9	5197	2	US-08-668-123-53	Sequence 53, Appli
C 479	12.4	68.9	5479	4	US-10-164-895-43	Sequence 43, Appli
C 480	12.4	68.9	5735	4	US-09-566-921-12	Sequence 12, Appli
C 481	12.4	68.9	7521	3	US-09-004-838-116	Sequence 116, App
C 482	12.4	68.9	7824	3	US-08-718-388-6	Sequence 6, Appli
C 483	12.4	68.9	12141	3	US-09-488-571-10	Sequence 10, Appli
C 484	12.4	68.9	12665	4	US-08-961-527-134	Sequence 134, App
C 485	12.4	68.9	16382	3	US-08-718-388-8	Sequence 8, Appli
C 486	12.4	68.9	16582	4	US-08-956-171E-53	Sequence 53, Appli
C 487	12.4	68.9	16592	4	US-08-781-986A-53	Sequence 53, Appli
C 488	12.4	68.9	25603	4	US-09-819-607-3	Sequence 3, Appli
C 489	12.4	68.9	31328	3	US-09-215-694-19	Sequence 19, Appli
C 490	12.4	68.9	37030	4	US-08-111-731A-25	Sequence 25, Appli
C 491	12.4	68.9	128779	4	US-09-497-855A-38	Sequence 38, Appli
C 492	12.4	68.9	259223	4	US-09-596-002-41	Sequence 41, Appli
C 493	12.4	68.9	536185	4	US-09-214-808-1	Sequence 1, Appli
C 494	12.2	67.8	19	4	US-09-422-978-9121	Sequence 9121, Ap
C 495	12.2	67.8	20	1	US-08-771-880A-51	Sequence 51, Appli
C 496	12.2	67.8	20	2	US-08-910-408-51	Sequence 51, Appli
C 497	12.2	67.8	20	3	US-09-249-215-51	Sequence 51, Appli
C 498	12.2	67.8	20	4	US-09-544-398B-323	Sequence 323, App
C 499	12.2	67.8	28	3	US-09-251-645-19	Sequence 19, Appli
C 500	12.2	67.8	31	4	US-09-603-052-10	Sequence 10, Appli

## ALIGNMENTS

```

RESULT 1
US-09-414-436-2/c
; Sequence 2, Application US/0941436
; Patent No. 6294384
; GENERAL INFORMATION:
; APPLICANT: Dell'Acqua, Giorgio
; APPLICANT: Mann, Michael J.
; TITLE OF INVENTION: Compositions and Methods Based Upon an Isoform of p53
; FILE REFERENCE: p53f
; CURRENT APPLICATION NUMBER: US/09/414,436
; EARLIER FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: 60/103,849
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1415
; ORGANISM: Rattus norvegicus
; US-09-414-436-2

Query Match      100.0%; Score 18; DB 3; Length 1415;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGACTGTGATCTCCAT 18
        |||
Db       41 CGACTGTGATCTCCAT 24

RESULT 2
US-08-184-009-214/c

```

Sequence 214, Application US/08184009  
Patent No. 5933975  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
APPLICANT: Tartaglia, James  
APPLICANT: Cox, William I.  
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 217  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/184,009  
FILING DATE: 19-JAN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2530  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066CURTMS  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-184-009-214

Query Match 91.1%; Score 16.4; DB 2; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18  
|||  
Db 27 CGACTGTGATCCTCCAT 10

RESULT 3  
US-08-458-356-214/c  
Sequence 214, Application US/08458356  
Patent No. 5942235  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
APPLICANT: Tartaglia, James  
APPLICANT: Cox, William I.  
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 217  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/458,356  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/184,009  
FILING DATE: 19-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2530  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066CURTMS  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-458-356-214

Query Match 91.1%; Score 16.4; DB 2; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18  
|||  
Db 27 CGACTGTGATCCTCCAT 10

RESULT 4  
US-08-796-101-45/c  
Sequence 45, Application US/08796101  
Patent No. 6183752  
GENERAL INFORMATION:  
APPLICANT: ERSTEIN, STEPHEN B.  
APPLICANT: FINKEL, TOREN  
APPLICANT: SPEIR, EDITH  
APPLICANT: ZHOU, YI FU  
APPLICANT: ZHU, JIANHUI  
APPLICANT: ERDILE, LORENE  
APPLICANT: PINCUS, STEVEN  
TITLE OF INVENTION: RESTENOSIS/ATHEROSCLEROSIS DIAGNOSIS,  
NUMBER OF SEQUENCES: 184  
TITLE OF INVENTION: PROPHYLAXIS AND THERAPY  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/796,101  
FILING DATE: 05-FEB-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: KOMALSKI, THOMAS J.  
REGISTRATION NUMBER: 32,147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 764-5574  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-796-101-45

Query Match 91.1%; Score 16.4; DB 3; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 5  
US-08-460-736-214/C  
Sequence 214, Application US/08460736  
Patent No. 6265189

GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
APPLICANT: Tartaglia, James  
APPLICANT: Cox, William I.  
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 217  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,736  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/184,009  
FILING DATE: 19-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2530  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066CURTMS  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-460-736-214

Query Match 91.1%; Score 16.4; DB 3; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 6  
US-09-535-370-214/C  
Sequence 214, Application US/09535370  
Patent No. 6537594

## GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo  
Tartaglia, James  
Cox, William I.  
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 217  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/535,370  
FILING DATE: 24-Mar-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/460,736  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2530  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066CURTMS  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 214:  
US-09-535-370-214

Query Match 91.1%; Score 16.4; DB 4; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 7  
US-09-663-667-214/C  
Sequence 214, Application US/09663667  
Patent No. 6780407

GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
Tartaglia, James  
Cox, William I.  
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 217  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/663,667  
FILING DATE: 15-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/184,009  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2530  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066CURTWS  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 214:  
US-09-663-667-214  
Query Match 91.1%; Score 16.4; DB 4; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCAT 18  
Db 27 CGACTGTGAATCCTCCAT 10  
RESULT 8  
US-08-585-593A-36/C  
Sequence 36, Application US/08585593A  
Patent No. 6503706  
GENERAL INFORMATION:  
APPLICANT: ABKEN, Hinrich J  
APPLICANT: ALBERT, Winfried  
APPLICANT: JUNGFER, Herbert  
TITLE OF INVENTION: METHOD OF IDENTIFYING HUMAN AND ANIMAL  
CELLS CAPABLE OF UNLIMITED PROLIFERATION OR TUMOR  
TITLE OF INVENTION: FORMATION  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/585,593A  
FILING DATE: 16-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94/02307  
FILING DATE: 13-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 43 23 727.4  
FILING DATE: 15-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Murray, Robert B.  
REGISTRATION NUMBER: 22,980  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-5000  
TELEFAX: (202) 638-4810  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-585-593A-36  
Query Match 85.6%; Score 15.4; DB 4; Length 396;  
Best Local Similarity 94.1%; Pred. No. 39;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GACTGTGAATCCTCCAT 18  
Db 302 GACTGTGAATCCTCCAT 286  
RESULT 9  
US-09-621-976-1600  
Sequence 1600, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jodert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent .pm  
SEQ ID NO 1600  
LENGTH: 527  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 136..435  
NAME/KEY: sig.peptide  
LOCATION: 136..279  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 5.80000019073486  
OTHER INFORMATION: seq ALIWAQVGVLA/GR  
US-09-621-976-1600  
Query Match 85.6%; Score 15.4; DB 4; Length 527;  
Best Local Similarity 94.1%; Pred. No. 40;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCA 17  
Db 430 CGACTGTGAATCCTCCA 446  
RESULT 10  
US-09-270-767-14686/C  
Sequence 14686, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 14686  
LENGTH: 2516  
TYPE: DNA  
ORGANISM: Drosophila melanogaster

US-09-270-767-14686

## Query Match

Best Local Similarity 85.6%; Score 15.4; DB 4; Length 2516;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGATCTCTCCA 17

Db 811 CGACTGTGATCTCTCCA 795

## RESULT 11

US-09-328-475C-5/c  
Sequence 5, Application US/09328475C  
Patent No. 6476207

## GENERAL INFORMATION:

APPLICANT: Zhang, Jimmy  
APPLICANT: Astel, Jon H.  
APPLICANT: Carroll III, Eddie  
APPLICANT: Endege, Wilson O.  
APPLICANT: Ford, Donna M.  
APPLICANT: Monahan, John E.  
APPLICANT: Schlegel, Robert  
APPLICANT: Steinmann, Kathleen E.  
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
CURRENT APPLICATION NUMBER: US/09/328,475C  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 5  
LENGTH: 1024  
TYPE: DNA  
ORGANISM: Homo Sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1024)  
OTHER INFORMATION: n = A,T,C or G  
US-09-328-475C-5

## Query Match

Best Local Similarity 83.3%; Score 15; DB 4; Length 1024;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ACTGTGATCTCTCCAT 18

Db 574 ACTGTGATCTCTCCAT 559

## RESULT 12

US-09-248-796A-8906/c  
Sequence 8906, Application US/09248796A  
Patent No. 6747137

## GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 8906  
LENGTH: 1107  
TYPE: DNA  
ORGANISM: Candida albicans  
US-09-248-796A-8906

## Query Match

82.2%; Score 14.8; DB 4; Length 1107;

Best Local Similarity 88.9%; Pred. No. 96;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGATCTCTCCAT 18

Db 283 CTACTGGATCTCTCCAT 266

## RESULT 13

US-09-621-976-18989/c  
Sequence 18989, Application US/09621976  
Patent No. 6639063

## GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 18989  
LENGTH: 478  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-18989

## Query Match

Best Local Similarity 80.0%; Score 14.4; DB 4; Length 478;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ACTGTGATCTCTCCAT 18

Db 115 ACTGTGATCTCTCCAT 100

## RESULT 14

US-09-186-276B-53/c  
Sequence 53, Application US/09186276B  
Patent No. 6388173

## GENERAL INFORMATION:

APPLICANT: Benfey, Philip  
APPLICANT: Dilaurenzio, Laura  
APPLICANT: Wysocka-Diller, Joanna  
APPLICANT: Malamy, Jocelyn E.  
APPLICANT: Pysh, Leonard  
APPLICANT: Helariutta, Yrjo  
TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses Thereof  
FILE REFERENCE: 5914-075-999  
CURRENT APPLICATION NUMBER: US/09/186,276B  
CURRENT FILING DATE: 1998-11-05  
PRIOR APPLICATION NUMBER: 08/842,445  
PRIOR FILING DATE: 1997-04-24  
PRIOR APPLICATION NUMBER: 08/638,617  
PRIOR FILING DATE: 1996-04-26  
NUMBER OF SEQ ID NOS: 79  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 53  
LENGTH: 1093  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-186-276B-53

## Query Match

Best Local Similarity 80.0%; Score 14.4; DB 3; Length 1093;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GACTGTGATCTCTCCA 17

Db 164 GACTGTGATCTCTCCA 149

RESULT 15  
US-08-842-445-53/C  
; Sequence 53, Application US/08842445A  
; Patent No. 6441270  
; GENERAL INFORMATION:  
; APPLICANT: Benfey et al.  
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 5914-056-999  
; CURRENT APPLICATION NUMBER: US/08/842,445A  
; CURRENT FILING DATE: 1997-04-24  
; EARLIER APPLICATION NUMBER: 08/638, 617  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 53  
; LENGTH: 1093  
; TYPE: DNA  
; ORGANISM: Plant  
US-08-842-445-53

Query Match 80.0%; Score 14.4; DB 4; Length 1093;  
Best Local Similarity 93.8%; Pred. No. 1.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17  
DB 164 GACTGTGAATCCTCCA 149

RESULT 16  
US-09-186-1888-53/C  
; Sequence 53, Application US/09186188B  
; Patent No. 6455672  
; GENERAL INFORMATION:  
; APPLICANT: Benfey et al.  
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 5914-074-999  
; CURRENT APPLICATION NUMBER: US/09/186,188B  
; CURRENT FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: 08/842,445  
; PRIOR FILING DATE: 1997-04-24  
; PRIOR APPLICATION NUMBER: 08/638, 617  
; PRIOR FILING DATE: 1996-04-26  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 53  
; LENGTH: 1093  
; TYPE: DNA  
; ORGANISM: Plant  
US-09-186-1888-53

Query Match 80.0%; Score 14.4; DB 4; Length 1093;  
Best Local Similarity 93.8%; Pred. No. 1.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17  
DB 164 GACTGTGAATCCTCCA 149

RESULT 17  
US-09-919-039-32/C  
; Sequence 32, Application US/09919039  
; Patent No. 6727066  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919, 039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113

PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 32  
; LENGTH: 1629  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6727066 3393861CB1  
US-09-919-039-32

Query Match 80.0%; Score 14.4; DB 4; Length 1629;  
Best Local Similarity 93.8%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACTGTGAATCCTCCAT 18  
DB 1074 ACAGTGAATCCTCCAT 1059

RESULT 18  
US-08-724-466B-3/C  
; Sequence 3, Application US/08724466B  
; Patent No. 6063606  
; GENERAL INFORMATION:  
; APPLICANT: Petkovich, P. Martin, White, Jay A.,  
; APPLICANT: Beckett, Barbara R., Jones, Glenville  
; TITLE OF INVENTION: Retinoid Metabolizing Protein  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Blake, Cassels & Graydon  
; STREET: Box 25, Commerce Court West  
; CITY: Toronto  
; ZIP: M5L 1A9  
; COUNTRY: Canada  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage  
; COMPUTER: COMPAQ, IBM PC compatible  
; OPERATING SYSTEM: MS-DOS 5.1  
; SOFTWARE: WORD PERFECT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,466B  
; FILING DATE: October 1, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/667,546  
; FILING DATE: June 21, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hunt, John C.  
; REGISTRATION NUMBER: 36,424  
; REFERENCE/DOCKET NUMBER: 50767/00004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 863-4344  
; TELEFAX: (416) 863-2653  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1850 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-724-466B-3

Query Match 80.0%; Score 14.4; DB 3; Length 1850;  
Best Local Similarity 93.8%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCC 16  
DB 1204 CGACTGTGAATCCTCC 1189

RESULT 19  
US-08-882-164D-3/C

Sequence 3, Application US/08882164D  
Patent No. 6306624  
GENERAL INFORMATION:  
APPLICANT: Petkovich, P. Martin, White, Jay A.,  
APPLICANT: Beckett, Barbara R., Jones, Glenville  
TITLE OF INVENTION: Retinoid Metabolizing Protein  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Blake, Cassels & Graydon  
STREET: Box 25, Commerce Court West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5L 1A9  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage  
COMPUTER: COMPAQ, IBM PC compatible  
OPERATING SYSTEM: MS-DOS 5.1  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,164D  
FILING DATE: June 25, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
APPLICATION NUMBER: 08/724,466  
FILING DATE: October 1, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1850 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-882-164D-3

Query Match 80.0%; Score 14.4; DB 3; Length 1850;  
Best Local Similarity 93.8%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCC 16  
Db 1204 CGACTGTGAATCCTCC 1189

RESULT 20  
US-09-399-081A-7  
Sequence 7, Application US/09399081A  
Patent No. 6646113  
GENERAL INFORMATION:  
APPLICANT: DREYFUS, Gideon  
APPLICANT: CHARROUX, Bernard  
APPLICANT: PELIZZONI, Livio  
APPLICANT: FISCHER, Utz  
APPLICANT: Liu, Qing  
TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS USEFUL FOR THE DIAGNOSIS AND TREAT  
FILE REFERENCE: 053893-5027  
CURRENT APPLICATION NUMBER: US/09/399,081A  
PRIOR APPLICATION NUMBER: 1999-09-17  
PRIOR FILING DATE: 1998-09-17  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patent version 3.1  
SEQ ID NO 7  
LENGTH: 3152  
TYPE: DNA

ORGANISM: Homo sapiens  
US-09-399-081A-7

Query Match 80.0%; Score 14.4; DB 4; Length 3152;  
Best Local Similarity 93.8%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17  
Db 2159 GACTGTGAATCCTCCA 2174

RESULT 21  
US-08-311-174-4/C  
Sequence 4, Application US/08311174  
Patent No. 5556776  
GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, MAKOTO  
APPLICANT: MIWA, KIYOSHI  
TITLE OF INVENTION: SUCRASE GENE DERIVED FROM CORYNEFORM  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,174  
FILING DATE: 23-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 046836/1992  
FILING DATE: 04-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5556776man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 10-699-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6911 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULAR TYPE: DNA (genomic)  
US-08-311-174-4

Query Match 80.0%; Score 14.4; DB 1; Length 6911;  
Best Local Similarity 93.8%; Pred. No. 2.1e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17  
Db 5815 GACTGTGAATCCTCCA 5800

RESULT 22  
US-08-206-176-3/C  
Sequence 3, Application US/08206176  
Patent No. 5639940  
GENERAL INFORMATION:



APPLICANT: Garner, Ian  
APPLICANT: Dalrymple, Michael A  
APPLICANT: Prunkard, Donna E  
APPLICANT: Foster, Donald C  
TITLE OF INVENTION: Production of Fibrinogen in Transgenic  
TITLE OF INVENTION: Animals  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Zymogenetics, Inc.  
STREET: 4225 Roosevelt Way, N.E.  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/206,176  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, Gary E  
REGISTRATION NUMBER: 31-648  
REFERENCE/DOCKET NUMBER: 93-15  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-547-8080 ext 322  
TELEFAX: 206-548-2329  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8878 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULAR TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: human fibrinogen B-beta chain  
FEATURE:  
NAME/KEY: misc RNA  
LOCATION: 1..469  
FEATURE:  
NAME/KEY: exon  
LOCATION: 470..583  
FEATURE:  
NAME/KEY: intron  
LOCATION: 584..3257  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3258..3449  
FEATURE:  
NAME/KEY: intron  
LOCATION: 3450..3938  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3939..4122  
FEATURE:  
NAME/KEY: intron  
LOCATION: 4123..5042  
FEATURE:  
NAME/KEY: exon  
LOCATION: 5043..5270  
FEATURE:  
NAME/KEY: intron  
LOCATION: 5271..5830  
FEATURE:  
NAME/KEY: exon  
LOCATION: 5831..5944  
FEATURE:  
NAME/KEY: intron  
LOCATION: 5945..6632  
FEATURE:

NAME/KEY: exon  
LOCATION: 6633..6758  
FEATURE:  
NAME/KEY: intron  
LOCATION: 6759..6966  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6967..7252  
FEATURE:  
NAME/KEY: intron  
LOCATION: 7253..7870  
FEATURE:  
NAME/KEY: exon  
LOCATION: 7871..8102  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 8103..8537  
FEATURE:  
NAME/KEY: misc RNA  
LOCATION: 8538..8878  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(470..583, 3258..3449, 3939..4122, 5043..5270,  
5831..5944, 6633..6758, 6967..7252, 7871..8102)  
US-08-206-176-3  
Query Match 80.0%; Score 14.4; DB 1; Length 8878;  
Best Local Similarity 93.8%; Pred. No. 2.2e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Cy 3 ACTGTGAATCTCCAT 18  
Db 7090 ACAGTGAATCTCCAT 7075  
RESULT 23  
US-09-614-221A-167/C  
Sequence 167, Application US/09614221A  
Patent No. 6723837  
GENERAL INFORMATION:  
APPLICANT: Karunanandaa, Balasubramanian  
APPLICANT: Yu, Jaehyuk  
APPLICANT: Kishore, Ganesh M.  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED  
WITH STEROL SYNTHESIS AND METABOLISM  
FILE REFERENCE: 16516.075  
CURRENT APPLICATION NUMBER: US/09/614,221A  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/142,981  
PRIOR FILING DATE: 1999-07-12  
NUMBER OF SEQ ID NOS: 626  
SEQ ID NO 167  
LENGTH: 1395  
TYPE: DNA  
ORGANISM: Saccharomyces cerevisiae  
US-09-614-221A-167  
Query Match 77.8%; Score 14; DB 4; Length 1395;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 5 TGTGAATCTCCAT 18  
Db 119 TGTGAATCTCCAT 106  
RESULT 24  
US-09-513-999C-36262  
Sequence 36262, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.

APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
Patent No. 6783961  
FILE REFERENCE: 59, US2, REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 36262  
LENGTH: 432  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 381  
OTHER INFORMATION: r=a or g  
US-09-513-999C-36262

Query Match 76.7%; Score 13.8; DB 4; Length 432;  
Best Local Similarity 88.2%; Pred. No. 2.9e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACTGTGAATCTCCAT 17  
DB 65 CCAGTGTAACTCCCA 81

RESULT 25  
US-09-389-681-417/c  
Sequence 417, Application US/09389681A  
Patent No. 6518237

GENERAL INFORMATION:  
APPLICANT: Yuqiu, Jiang  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.470C3  
CURRENT APPLICATION NUMBER: US/09/389,681A  
CURRENT FILING DATE: 1999-09-02  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 417  
LENGTH: 514  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-389-681-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;  
Best Local Similarity 88.2%; Pred. No. 3e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18  
DB 338 GACTGTGAATCTCCAT 322

RESULT 26  
US-09-620-405B-417/c  
Sequence 417, Application US/09620405B  
Patent No. 6528054

GENERAL INFORMATION:  
APPLICANT: Jiang, Yuqiu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hedler, William T.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.470C8  
CURRENT APPLICATION NUMBER: US/09/620,405B  
CURRENT FILING DATE: 2000-07-20  
NUMBER OF SEQ ID NOS: 495  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 417  
LENGTH: 514  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-620-405B-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;  
Best Local Similarity 88.2%; Pred. No. 3e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18  
DB 338 GACTGTGAATCTCCAT 322

RESULT 27  
US-09-433-826B-417/c  
Sequence 417, Application US/09433826B  
Patent No. 6579973

GENERAL INFORMATION:  
APPLICANT: Jiang, Yuqiu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.470C4  
CURRENT APPLICATION NUMBER: US/09/433,826B  
CURRENT FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 474  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 417  
LENGTH: 514  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-433-826B-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;  
Best Local Similarity 88.2%; Pred. No. 3e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18  
DB 338 GACTGTGAATCTCCAT 322

RESULT 28  
US-09-604-287A-417/c  
Sequence 417, Application US/09604287A  
Patent No. 6586572

GENERAL INFORMATION:  
APPLICANT: Jiang, Yuqiu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hedler, William T.  
TITLE OF INVENTION: COMPOSITIONS OF BREAST CANCER  
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.470C7  
CURRENT APPLICATION NUMBER: US/09/604,287A  
CURRENT FILING DATE: 2000-06-22  
NUMBER OF SEQ ID NOS: 489  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 417  
LENGTH: 514  
TYPE: DNA

ORGANISM: Homo sapiens  
US-09-604-287A-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
||| |||||  
Db 338 GACTGTGAATCCTCCAT 322

RESULT 29  
US-09-834-759-417/c  
; Sequence 417, Application US/09834759  
; Patent No. 6680197  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jianshun  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.470C9  
; CURRENT APPLICATION NUMBER: US/09/834,759  
; CURRENT FILING DATE: 2001-04-13  
; NUMBER OF SEQ ID NOS: 547  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 417  
; LENGTH: 514  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-834-759-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
||| |||||  
Db 338 GACTGTGAATCCTCCAT 322

RESULT 30  
US-09-590-751A-417/c  
; Sequence 417, Application US/09590751A  
; Patent No. 6756477  
; GENERAL INFORMATION:  
; APPLICANT: Yugu, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jianshun  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.470C6  
; CURRENT APPLICATION NUMBER: US/09/590,751A  
; CURRENT FILING DATE: 2000-06-08  
; NUMBER OF SEQ ID NOS: 479  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 417  
; LENGTH: 514  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-590-751A-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
||| |||||  
Db 338 GACTGTGAATCCTCCAT 322

RESULT 31  
US-09-489-039A-2781  
; Sequence 2781, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 2781  
; LENGTH: 516  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-2781

Query Match 76.7%; Score 13.8; DB 4; Length 516;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACTGTGAATCCTCCA 17  
||| |||||  
Db 198 GACTGTGAATCCTCCA 214

RESULT 32  
US-09-389-681-258/c  
; Sequence 258, Application US/09389681A  
; Patent No. 6518237  
; GENERAL INFORMATION:  
; APPLICANT: Yugu, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jianshun  
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.470C3  
; CURRENT APPLICATION NUMBER: US/09/389,681A  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 258  
; LENGTH: 519  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-389-681-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
||| |||||  
Db 343 GACTGTGAATCCTCCAT 327

RESULT 33  
US-09-620-405B-258/c  
; Sequence 258, Application US/09620405B  
; Patent No. 6528054  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.

APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C8  
CURRENT APPLICATION NUMBER: US/09/620,405B  
CURRENT FILING DATE: 2000-07-20  
NUMBER OF SEQ ID NOS: 495  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 258  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-620-405B-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
DB 343 GACTTGAATCCTCCAT 327

RESULT 34  
US-09-339-338-258/C  
Sequence 258, Application US/09339338A  
Patent No. 6573368  
GENERAL INFORMATION:  
APPLICANT: Yugu, Jiang  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
FILE REFERENCE: 210121.470C2  
CURRENT APPLICATION NUMBER: US/09/339,338A  
CURRENT FILING DATE: 1999-06-23  
NUMBER OF SEQ ID NOS: 315  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 258  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-339-338-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
DB 343 GACTTGAATCCTCCAT 327

RESULT 35  
US-09-433-826B-258/C  
Sequence 258, Application US/09433826B  
Patent No. 6579973  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
FILE REFERENCE: 210121.470C4  
CURRENT APPLICATION NUMBER: US/09/433,826B  
CURRENT FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 474  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 258  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-433-826B-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
DB 343 GACTTGAATCCTCCAT 327

RESULT 36  
US-09-604-287A-258/C  
Sequence 258, Application US/09604287A  
Patent No. 6586572  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C7  
CURRENT APPLICATION NUMBER: US/09/604,287A  
CURRENT FILING DATE: 2000-06-22  
NUMBER OF SEQ ID NOS: 489  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 258  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-604-287A-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
DB 343 GACTTGAATCCTCCAT 327

RESULT 37  
US-09-834-759-258/C  
Sequence 258, Application US/09834759  
Patent No. 6680197  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C9  
CURRENT APPLICATION NUMBER: US/09/834,759  
CURRENT FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 547  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 258  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-834-759-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCCTCCAT 18  
|||  
Db 343 GACTGTGAATCCTCCAT 327

RESULT 38  
US-09-590-751A-258/c  
; Sequence 258, Application US/09590751A  
; Patent No. 6756477

GENERAL INFORMATION:  
; APPLICANT: Yuqin, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.470C6  
; CURRENT APPLICATION NUMBER: US/09/590,751A  
; CURRENT FILING DATE: 2000-06-08  
; NUMBER OF SEQ ID NOS: 479  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 258  
; LENGTH: 519  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-590-751A-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;  
Best Local Similarity 88.2%; Pred. No. 3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCCTCCAT 18  
|||  
Db 343 GACTGTGAATCCTCCAT 327

RESULT 39  
US-09-270-767-11839/c  
; Sequence 11839, Application US/09270767  
; Patent No. 6703491  
GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11839  
; LENGTH: 598  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-11839

Query Match 76.7%; Score 13.8; DB 4; Length 598;  
Best Local Similarity 88.2%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCA 17  
|||  
Db 348 CGCAGTGAATCCTCCA 332

RESULT 40  
US-09-371-307-79/c  
; Sequence 79, Application US/09371307A  
; Patent No. 6723897  
; GENERAL INFORMATION:

APPLICANT: Brown, Sherri M.  
APPLICANT: Heck, Gregory R.  
APPLICANT: Piller, Kenneth J.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Ellich, Ted D.  
APPLICANT: Logusch, Eugene W.  
APPLICANT: Rao, Sudabathula  
APPLICANT: Ream, Joel E.  
APPLICANT: Logusch, Sherry J.  
TITLE OF INVENTION: Methods for controlling gibberellin levels  
FILE REFERENCE: MOST:216  
CURRENT APPLICATION NUMBER: US/09/371,307A  
CURRENT FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 79  
LENGTH: 1041  
TYPE: DNA  
ORGANISM: Cucurbita maxima  
US-09-371-307-79

Query Match 76.7%; Score 13.8; DB 4; Length 1041;  
Best Local Similarity 88.2%; Pred. No. 3.3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCCTCCAT 18  
|||  
Db 396 GACTGTGAATCCTCCAT 380

RESULT 41  
US-09-257-580-1/c  
; Sequence 1, Application US/09257580  
; Patent No. 6307036  
GENERAL INFORMATION:  
APPLICANT: Yorkshire Cancer Research  
TITLE OF INVENTION: Tumour Suppressor Gene  
FILE REFERENCE: Canine p53  
CURRENT APPLICATION NUMBER: US/09/257,580  
CURRENT FILING DATE: 1999-02-25  
PRIOR APPLICATION NUMBER: 9804178.3  
PRIOR FILING DATE: 1998-02-28  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1146  
TYPE: DNA  
ORGANISM: Canis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1146)  
US-09-257-580-1

Query Match 76.7%; Score 13.8; DB 3; Length 1146;  
Best Local Similarity 88.2%; Pred. No. 3.4e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCCTCCAT 18  
|||  
Db 17 GACTGTGAATCCTCCAT 1

RESULT 42  
US-09-257-580-4/c  
; Sequence 4, Application US/09257580  
; Patent No. 6307036  
GENERAL INFORMATION:  
APPLICANT: Yorkshire Cancer Research  
TITLE OF INVENTION: Tumour Suppressor Gene  
FILE REFERENCE: Canine p53  
CURRENT APPLICATION NUMBER: US/09/257,580  
CURRENT FILING DATE: 1999-02-25  
PRIOR APPLICATION NUMBER: 9804178.3

PRIOR FILING DATE: 1998-02-28  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 1146  
TYPE: RNA  
ORGANISM: canis  
US-09-257-580-4

Query Match 76.7%; Score 13.8; DB 3; Length 1146;  
Best Local Similarity 88.2%; Pred. No. 3.4e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
DB 17 GACTGTGAATCCTCCAT 1

RESULT 43  
US-09-673-395A-57/c  
Sequence 57, Application US/09673395A  
Patent No. 6620923  
GENERAL INFORMATION:  
APPLICANT: SPECHT, THOMAS  
APPLICANT: HINZMANN, BERND  
APPLICANT: SCHMITT, ARMIN  
APPLICANT: PILARSKY, CHRISTIAN  
APPLICANT: DAHL, EDGAR  
APPLICANT: ROSENTHAL, ANDRE  
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE  
FILE REFERENCE: ALBRE-12  
CURRENT APPLICATION NUMBER: US/09/673,395A  
CURRENT FILING DATE: 2000-10-17  
NUMBER OF SEQ ID NOS: 637  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 57  
LENGTH: 1166  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-673-395A-57

Query Match 76.7%; Score 13.8; DB 4; Length 1166;  
Best Local Similarity 88.2%; Pred. No. 3.4e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
DB 1128 GACTGTGAATCCTCCAT 1112

RESULT 44  
US-09-270-767-12871/c  
Sequence 12871, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12871  
LENGTH: 1270  
TYPE: DNA  
ORGANISM: Drosophila melanogaster  
US-09-270-767-12871

Query Match 76.7%; Score 13.8; DB 4; Length 1270;  
Best Local Similarity 88.2%; Pred. No. 3.4e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18

DB 686 GTCTGTGAATCCTCCAT 670

RESULT 45  
US-10-000-489-87/c  
Sequence 87, Application US/10000489  
Patent No. 6794363  
GENERAL INFORMATION:  
APPLICANT: Benjamin, Stephane  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: 91.US6.DIV  
CURRENT APPLICATION NUMBER: US/10/000,489  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: Jpatent  
SEQ ID NO 87  
LENGTH: 1431  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..151  
NAME/KEY: CDS  
LOCATION: 152..655  
NAME/KEY: 3'UTR  
LOCATION: 656..1431  
NAME/KEY: polyA\_signal  
LOCATION: 1399..1404  
NAME/KEY: polyA\_site  
LOCATION: 1416..1431  
US-10-000-489-87

Query Match 76.7%; Score 13.8; DB 4; Length 1431;  
Best Local Similarity 88.2%; Pred. No. 3.5e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
DB 1109 GACTGTGAATCCTCCAT 1093

RESULT 46  
US-10-000-489-89/c  
Sequence 89, Application US/10000489  
Patent No. 6794363  
GENERAL INFORMATION:  
APPLICANT: Benjamin, Stephane  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: 91.US6.DIV  
CURRENT APPLICATION NUMBER: US/10/000,489  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277

PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: JPatent  
SEQ ID NO 89  
LENGTH: 1431  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..151  
NAME/KEY: CDS  
LOCATION: 152..655  
NAME/KEY: 3'UTR  
LOCATION: 656..1431  
NAME/KEY: polyA\_signal  
LOCATION: 1399..1404  
NAME/KEY: polyA\_site  
LOCATION: 1416..1431  
US-10-000-489-89

Query Match 76.7%; Score 13.8; DB 4; Length 1431;  
Best Local Similarity 88.2%; Pred. No. 3.5e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
Db 1109 GACTGTGAATCCTCCAT 1093

RESULT 47  
US-09-248-796A-5727  
Sequence 5727, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 5727  
LENGTH: 1458  
TYPE: DNA  
ORGANISM: Candida albicans  
US-09-248-796A-5727

Query Match 76.7%; Score 13.8; DB 4; Length 1458;  
Best Local Similarity 88.2%; Pred. No. 3.5e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
Db 631 GACTGTGAATCCTCCAT 647

RESULT 48  
US-09-719-919A-2/c  
Sequence 2, Application US/09719919A  
Patent No. 6680180  
GENERAL INFORMATION:  
APPLICANT: Jegia, Timothy James  
APPLICANT: ICAGEN Incorporated  
TITLE OF INVENTION: Kv6.2, a Voltage-Gated Potassium Channel Subunit  
FILE REFERENCE: 018512-001410US

CURRENT APPLICATION NUMBER: US/09/719,919A  
CURRENT FILING DATE: 2001-02-22  
PRIOR APPLICATION NUMBER: US 60/091,466  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: WO PCT/US99/14945  
PRIOR FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1518  
TYPE: DNA  
ORGANISM: Mus sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1518)  
OTHER INFORMATION: mouse alpha subunit of heteromeric voltage-gated  
OTHER INFORMATION: potassium channel Kv6.2  
US-09-719-919A-2

Query Match 76.7%; Score 13.8; DB 4; Length 1518;  
Best Local Similarity 88.2%; Pred. No. 3.5e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17  
Db 642 CGACTGTGAGTCTCCA 626

RESULT 49  
US-08-687-590-60  
Sequence 60, Application US/08687590  
Patent No. 6255070  
GENERAL INFORMATION:  
APPLICANT: Willison, Keith Robert  
APPLICANT: Kubota, Hiroshi  
APPLICANT: Ashworth, Alan  
TITLE OF INVENTION: Folding Proteins  
NUMBER OF SEQUENCES: 74  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,590  
FILING DATE: 31-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/00192  
FILING DATE: 31-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9401791.0  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9418234.2  
FILING DATE: 09-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Baastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 084619-000000US  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1734 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

```

;      FEATURE:
;      NAME/KEY:      CDS
;      LOCATION:      3..1637
US-08-687-590-60

```

Query Match	76.7%	Score 13.8	DB 3	Length 1734
Best Local Similarity	88.2%	Pred. No. 3.6e+02		
Matches 15, Conservative	0	Mismatches 2	Indels 0	Gaps 0

Qy	1	CGACTGTGAATCCTCCA	17
Db	1576	CCACTGTGATCCTCCA	1592

```

, RESULT 50
, US-09-673-395A-140
, Sequence 140 Application US/09673395A
, Patent No. 6620923
, GENERAL INFORMATION:
, APPLICANT: SPECHT, THOMAS
, APPLICANT: HINZMANN, BERND
, APPLICANT: SCHMITT, ARMIN
, APPLICANT: PILARSKY, CHRISTIAN
, APPLICANT: DAHL, EDGAR
, APPLICANT: ROSENTHAL, ANDRE
, TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
, FILE REFERENCE: ALBRE-12
, CURRENT APPLICATION NUMBER: US/09/673,395A
, CURRENT FILING DATE: 2000-10-17
, NUMBER OF SEQ ID NOS: 637
, SOFTWARE: PatentIn Ver. 2.1
, SEQ ID NO 140
, LENGTH: 1938
, TYPE: DNA
, ORGANISM: Homo sapiens
, US-09-673-395A-140

```

Query Match	76.7%	Score 13.8;	DB 4;	Length 1938;
Best Local Similarity	88.2%	Pred. No. 3.6e+02;		
Matches 15;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```

Qy      2 GACTGTGAATCCTCCAT 18
         |||||
Db      1621 GACTGTGAATCCTTGAT 1637

```

Search completed: January 14, 2005, 17:39:40  
Job time : 99.9474 Secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:35:09 ; Search time 2369.37 Seconds  
(without alignments)  
276.831 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18

Sequence: 1 CGACTGATCCTCCAT 18

Scoring table: IDENTITY NUC

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 500 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gss1: \*  
9: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	237	2	BB571570 BB571570
2	18	100.0	280	2	BB548882 UT-R-A0-a
3	18	100.0	302	2	BB567223 BB567223
4	18	100.0	319	2	BB126959 DEPA0708
5	18	100.0	376	6	CB691142 AMGNNUC:C
6	18	100.0	407	2	BB127536 DEPA1285
7	18	100.0	427	6	CB795162 AMGNNUC:S
8	18	100.0	429	6	CB372427 UT-R-G00-
9	18	100.0	471	6	CB732629 AMGNNUC:M
10	18	100.0	479	2	BB128412 DEPA2162
11	18	100.0	559	2	BB54812 UT-R-E0-C
12	18	100.0	559	6	CB609941 AMGNNUC:U
13	18	100.0	562	7	CP978067 PAR134.R
14	18	100.0	589	6	CB583399 AMGNNUC:N
15	18	100.0	715	7	CC040016 AGENCOURT
16	18	100.0	744	7	CC0394200 AGENCOURT
17	18	100.0	792	7	CC482053 AGENCOURT
18	18	100.0	797	7	CC477556 AGENCOURT
19	18	100.0	802	7	CC0560520 AGENCOURT
20	18	100.0	845	6	CB315809 AGENCOURT
21	18	100.0	846	6	CB602723 AGENCOURT
22	17	94.4	537	5	BB896918 X047G05.P
23	17	94.4	717	6	CB829200 BN40.041G
24	17	94.4	726	8	BB186569 CH230-340

25	17	94.4	1075	9	CNS0591U
26	16.4	91.1	175	7	CK34427 CK34427
27	16.4	91.1	188	7	CN685546
28	16.4	91.1	223	2	BB585084
29	16.4	91.1	245	2	BB604900
30	16.4	91.1	252	7	CN674653
31	16.4	91.1	266	4	BG062707
32	16.4	91.1	282	9	CG668669
33	16.4	91.1	332	5	BY232607
34	16.4	91.1	340	5	BY021968
35	16.4	91.1	351	5	BY010085
36	16.4	91.1	354	5	BY192197
37	16.4	91.1	354	7	D76778
38	16.4	91.1	397	5	BY043120
39	16.4	91.1	407	7	CK577863
40	16.4	91.1	428	2	AM209975
41	16.4	91.1	430	2	BB846564
42	16.4	91.1	447	6	CAB95446
43	16.4	91.1	447	9	CG501482
44	16.4	91.1	462	2	BB588888
45	16.4	91.1	466	7	CP726263
46	16.4	91.1	486	6	CA533945
47	16.4	91.1	506	6	CA565317
48	16.4	91.1	528	9	CG593379
49	16.4	91.1	529	6	CD560905
50	16.4	91.1	530	6	CA533770
51	16.4	91.1	545	7	CP892817
52	16.4	91.1	551	9	CG568015
53	16.4	91.1	555	6	CF162401
54	16.4	91.1	562	4	B1525965
55	16.4	91.1	572	5	BQ266772
56	16.4	91.1	575	7	CN683981
57	16.4	91.1	585	9	CG508476
58	16.4	91.1	596	1	A1731576
59	16.4	91.1	600	4	B1985246
60	16.4	91.1	600	5	BU923192
61	16.4	91.1	608	7	CK577814
62	16.4	91.1	608	7	CK577851
63	16.4	91.1	610	7	CP897981
64	16.4	91.1	617	1	AA030090
65	16.4	91.1	622	2	BB660925
66	16.4	91.1	622	6	BY731658
67	16.4	91.1	627	2	BB380449
68	16.4	91.1	634	4	B1150716
69	16.4	91.1	638	2	BB654885
70	16.4	91.1	643	2	BB663747
71	16.4	91.1	671	2	BB634494
72	16.4	91.1	692	2	AM476667
73	16.4	91.1	695	4	BJ817162
74	16.4	91.1	708	6	CG963753
75	16.4	91.1	713	6	CB948431
76	16.4	91.1	745	2	BB916329
77	16.4	91.1	758	2	BB581124
78	16.4	91.1	764	1	AV751729
79	16.4	91.1	768	4	BG864885
80	16.4	91.1	775	4	B1105128
81	16.4	91.1	785	4	B1652413
82	16.4	91.1	821	6	CB566296
83	16.4	91.1	846	5	BO953341
84	16.4	91.1	872	4	B1416028
85	16.4	91.1	881	2	B1616854
86	16.4	91.1	882	6	CB203038
87	16.4	91.1	904	5	BU517804
88	16.4	91.1	905	5	BU147811
89	16.4	91.1	911	2	BF133788
90	16.4	91.1	915	5	BQ900722
91	16.4	91.1	923	5	BO928343
92	16.4	91.1	923	5	BO965113
93	16.4	91.1	932	5	BQ928016
94	16.4	91.1	936	4	BG969702
95	16.4	91.1	938	4	BG293270
96	16.4	91.1	939	6	CB204419
97	16.4	91.1	954	2	BB534082

AL326667	Tetradon
CK334427	H3026D07-
CN685546	E021008-
BB585084	BB585084
BB604900	BB604900
CN674653	A0951H02-
BG062707	L0956G04-
CG668669	OST464816
BY232607	BY232607
BY021968	BY021968
BY010085	BY010085
BY192197	BY192197
D76778	MU74C11 mo
BY043120	BY043120
CK577863	IST_WT5_1
AM209975	u15d12.y
BB846564	BB846564
CAB95446	B0192B03-
CG501482	OST45353
BB588888	BB588888
CP726263	UT-M-G20-
CA533945	C0409B09-
CA565317	K0336E11-
CG593379	OST250499
CD560905	B0421A08-
CA533770	C0406D02-
CP892817	A0115B04-
CG568015	OST194703
CF162401	B0711A07-
B1525965	602926887
BQ266772	NISC_F414
CN683981	E0186D04-
CG508476	OST959106
A1731576	BN7GH101
B1985246	3133-29 M
BU923192	7051-51 M
CK577814	IST_WT5_1
CK577851	IST_WT5_1
CP897981	A0232B10-
AA030090	m127d03.x
BB660925	BB660925
BY731658	BY731658
BB380449	601270523
B1150716	602914801
BB654885	BB654885
BB663747	BB663747
BB634494	BB634494
AM476667	uq79h11.y
BJ817162	BJ817162
CG963753	602831395
CB948431	AGENCOURT
BB916329	601666905
BB581124	602100258
AV751729	AV751729
BG864885	602799329
B1105128	602893382
B1652413	603299691
CB566296	AGENCOURT
BO953341	AGENCOURT
B1416028	602987181
B1616854	601774885
CB203038	AGENCOURT
BU517804	AGENCOURT
BU147811	AGENCOURT
BF133788	601778707
BQ900722	AGENCOURT
BO928343	AGENCOURT
BO965113	AGENCOURT
BQ928016	AGENCOURT
BG969702	602837471
BG293270	602390848
CB204419	AGENCOURT
BB534082	602075070

C 98	16.4	91.1	954	5	BUS12162	BUS12162	AGENCOURT	C 171	15.4	85.6	424	2	AM184013	AM184013	xj390f03.x
C 99	16.4	91.1	970	5	BO891359	BO891359	AGENCOURT	C 172	15.4	85.6	427	5	BP624231	BP624231	
C 100	16.4	91.1	982	5	BO958441	BO958441	AGENCOURT	C 173	15.4	85.6	429	5	BT388889	BT388889	949047A08
C 101	16.4	91.1	992	5	BO963167	BO963167	AGENCOURT	C 174	15.4	85.6	429	5	BP634296	BP634296	
C 102	16.4	91.1	1053	3	AK079711	AK079711	Mus muscu	C 175	15.4	85.6	429	6	CF134938	CF134938	
C 103	16.4	91.1	1093	3	BQ887132	BQ887132	AGENCOURT	C 176	15.4	85.6	433	6	CA043074	CA043074	BBAlmgfCB0
C 104	16.4	91.1	1139	2	BF795142	BF795142	602256480	C 177	15.4	85.6	443	8	BZ857345	BZ857345	CH240.232
C 105	16.4	91.1	1159	3	AK086845	AK086845	Mus muscu	C 178	15.4	85.6	448	7	H18273	H18273	YPA6807.r1
C 106	16.4	91.1	2901	3	AK087834	AK087834	Mus muscu	C 179	15.4	85.6	450	8	BZ984603	BZ984603	PUGHOT1TD
C 107	16.4	91.1	3276	3	AK083750	AK083750	Mus muscu	C 180	15.4	85.6	453	4	BT243070	BT243070	
C 108	16	88.9	171	1	AL610480	AL610480	US010460	C 181	15.4	85.6	464	7	CN070806	CN070806	1021003E0
C 109	16	88.9	222	8	AZ345159	AZ345159	1M0079N02	C 182	15.4	85.6	467	4	BT243925	BT243925	
C 110	16	88.9	440	9	CG675515	CG675515	1M0079N02	C 183	15.4	85.6	468	4	BG599937	BG599937	EST504832
C 111	16	88.9	488	1	AL816299	AL816299	AL816299	C 184	15.4	85.6	469	7	CK696304	CK696304	ZFT01-P00
C 112	16	88.9	503	8	A0536049	A0536049	RPCT-11-3	C 185	15.4	85.6	475	1	AV557161	AV557161	AV557161
C 113	16	88.9	551	8	AO143786	AO143786	HS_3075_B	C 186	15.4	85.6	477	5	BQ339673	BQ339673	
C 114	16	88.9	558	8	CG672956	CG672956	1E80704-T	C 187	15.4	85.6	477	6	CD058843	CD058843	3529_1_10
C 115	16	88.9	563	8	A0536391	A0536391	RPCT-11-3	C 188	15.4	85.6	481	1	AA440442	AA440442	LD15288.5
C 116	16	88.9	1041	8	CC238336	CC238336	CH261-118	C 189	15.4	85.6	485	8	AQ247352	AQ247352	HS_2059_B
C 117	16	88.9	1057	4	BM923125	BM923125	AGENCOURT	C 190	15.4	85.6	488	5	BO508499	BO508499	EST615914
C 118	16	88.9	1160	8	CC273218	CC273218	CH261-142	C 191	15.4	85.6	489	1	AA846850	AA846850	342404.8
C 119	15.4	85.6	148	4	BG983069	BG983069	PM1-CN009	C 192	15.4	85.6	489	1	AA951763	AA951763	
C 120	15.4	85.6	154	7	CN402428	CN402428	170005321	C 193	15.4	85.6	489	4	BT700255	BT700255	
C 121	15.4	85.6	169	2	BF946959	BF946959	MRO-NN019	C 194	15.4	85.6	490	2	BF945442	BF945442	
C 122	15.4	85.6	202	1	AV211510	AV211510	AV211510	C 195	15.4	85.6	491	2	BF945416	BF945416	
C 123	15.4	85.6	206	2	BBS90763	BBS90763	BBS90763	C 196	15.4	85.6	495	5	BX508749	BX508749	DREPE8861
C 124	15.4	85.6	213	7	Z38569	Z38569	HSC0C022_P	C 197	15.4	85.6	499	5	BO339220	BO339220	
C 125	15.4	85.6	224	4	BT050299	BT050299	PM2-CN001	C 198	15.4	85.6	499	9	CR221862	CR221862	Forward_B
C 126	15.4	85.6	235	2	BP958492	BP958492	QV2-NN004	C 199	15.4	85.6	503	2	BB754997	BB754997	
C 127	15.4	85.6	238	1	AV208003	AV208003	AV208003	C 200	15.4	85.6	506	2	BF409148	BF409148	UT-R-BT1-
C 128	15.4	85.6	245	9	CG192419	CG192419	POTAH86TB	C 201	15.4	85.6	506	2	BF945429	BF945429	MRO-NN019
C 129	15.4	85.6	250	2	BB090142	BB090142	BB090142	C 202	15.4	85.6	509	1	AI774453	AI774453	EST25553
C 130	15.4	85.6	254	1	AV284225	AV284225	AV284225	C 203	15.4	85.6	511	1	AI520014	AI520014	LD39934.5
C 131	15.4	85.6	255	1	AV215431	AV215431	AV215431	C 204	15.4	85.6	512	7	CN402431	CN402431	
C 132	15.4	85.6	262	1	AV133362	AV133362	AV133362	C 205	15.4	85.6	515	6	CB958705	CB958705	AGENCOURT
C 133	15.4	85.6	268	1	AV168774	AV168774	AV168774	C 206	15.4	85.6	519	2	BE512332	BE512332	946069C06
C 134	15.4	85.6	298	4	BI430863	BI430863	949062B04	C 207	15.4	85.6	524	7	CO516823	CO516823	61366998B
C 135	15.4	85.6	298	6	CP040507	CP040507	QCI14607.	C 208	15.4	85.6	525	4	BM686378	BM686378	UT-E-CRO-
C 136	15.4	85.6	303	4	BG983689	BG983689	PM1-CN015	C 209	15.4	85.6	527	8	AQ040939	AQ040939	HS_5048_B
C 137	15.4	85.6	308	1	AI665168	AI665168	605007F08	C 210	15.4	85.6	528	4	BM684111	BM684111	
C 138	15.4	85.6	322	2	BE149530	BE149530	RC1-RT025	C 211	15.4	85.6	532	7	CN402426	CN402426	170006002
C 139	15.4	85.6	329	2	AM281052	AM281052	AM281052	C 212	15.4	85.6	532	8	AQ069722	AQ069722	HS_5519_B
C 140	15.4	85.6	332	7	CG670213	CG670213	OST469229	C 213	15.4	85.6	534	4	BG489781	BG489781	602518847
C 141	15.4	85.6	341	9	CN402422	CN402422	170004241	C 214	15.4	85.6	540	1	AI519195	AI519195	LD38706.5
C 142	15.4	85.6	345	4	BI396316	BI396316	949047A08	C 215	15.4	85.6	540	4	BM64983	BM64983	
C 143	15.4	85.6	354	4	BI204750	BI204750	EST522790	C 216	15.4	85.6	541	1	AI739822	AI739822	
C 144	15.4	85.6	357	5	BO505461	BO505461	EST612876	C 217	15.4	85.6	546	8	BH775922	BH775922	
C 145	15.4	85.6	360	9	CC794143	CC794143	SAUK_0436	C 218	15.4	85.6	547	4	BI679635	BI679635	
C 146	15.4	85.6	361	5	BQ108514	BQ108514	shpID000	C 219	15.4	85.6	548	4	BI214393	BI214393	
C 147	15.4	85.6	371	5	BY416963	BY416963	BY416963	C 220	15.4	85.6	549	1	AA246762	AA246762	
C 148	15.4	85.6	372	8	AQ068980	AQ068980	HS_2255_B	C 221	15.4	85.6	549	4	BM822510	BM822510	
C 149	15.4	85.6	373	5	BY403891	BY403891	BY403891	C 222	15.4	85.6	550	1	AI668485	AI668485	
C 150	15.4	85.6	375	8	AQ072605	AQ072605	HS_2247_A	C 223	15.4	85.6	550	2	BE509950	BE509950	946069C06
C 151	15.4	85.6	378	4	BI417161	BI417161	949053H12	C 224	15.4	85.6	550	5	BQ513487	BQ513487	
C 152	15.4	85.6	378	4	BI674193	BI674193	949053H12	C 225	15.4	85.6	551	9	CG347657	CG347657	CO1C087TV
C 153	15.4	85.6	383	1	AL927205	AL927205	AL927205	C 226	15.4	85.6	551	4	BT566119	BT566119	
C 154	15.4	85.6	383	2	BF804230	BF804230	115-CI014	C 227	15.4	85.6	554	4	BI167432	BI167432	
C 155	15.4	85.6	385	5	BU667874	BU667874	MC01016E1	C 228	15.4	85.6	559	4	BI396397	BI396397	949063B04
C 156	15.4	85.6	387	4	BI358975	BI358975	949053H12	C 229	15.4	85.6	560	5	BT786357	BT786357	
C 157	15.4	85.6	388	6	CA521086	CA521086	KSI1020H1	C 230	15.4	85.6	561	2	BE639387	BE639387	
C 158	15.4	85.6	391	1	AA419137	AA419137	vz34B09.r	C 231	15.4	85.6	562	4	BI485698	BI485698	
C 159	15.4	85.6	392	2	BF591538	BF591538	nacl15A08.	C 232	15.4	85.6	564	8	AZ343732	AZ343732	1M0077F19
C 160	15.4	85.6	399	6	BY634653	BY634653	BY634653	C 233	15.4	85.6	565	1	AI533840	AI533840	SD05810.5
C 161	15.4	85.6	399	6	BY668011	BY668011	BY668011	C 234	15.4	85.6	566	2	AM218991	AM218991	EST301463
C 162	15.4	85.6	404	5	BP658184	BP658184	BP658184	C 235	15.4	85.6	566	7	CN402424	CN402424	170005325
C 163	15.4	85.6	404	5	BP647985	BP647985	BP647985	C 236	15.4	85.6	566	7	CO169279	CO169279	NDL1_5_HO
C 164	15.4	85.6	405	5	BU667476	BU667476	MC01006E0	C 237	15.4	85.6	569	6	CB070051	CB070051	1625B06.y
C 165	15.4	85.6	405	5	BP662842	BP662842	BP662842	C 238	15.4	85.6	571	1	AI712140	AI712140	605063G06
C 166	15.4	85.6	413	1	AV681725	AV681725	AV681725	C 239	15.4	85.6	574	1	AI518656	AI518656	
C 167	15.4	85.6	417	5	BY402841	BY402841	BY402841	C 240	15.4	85.6	575	2	BF520360	BF520360	EST457830
C 168	15.4	85.6	417	6	CF232938	CF232938	PAUX0001	C 241	15.4	85.6	577	4	BM647900	BM647900	K-EST0127
C 169	15.4	85.6	419	5	BY395750	BY395750	BY395750	C 242	15.4	85.6	577	6	CA842369	CA842369	1t26D10.y
C 170	15.4	85.6	420	8	AZ958129	AZ958129	2M0225114	C 243	15.4	85.6	578	1	AI519180	AI519180	LD38672.5

C 244	15.4	85.6	578	6	CD059078	3529	1	10	CD059078	3529	1	10
245	15.4	85.6	578	9	CC717411	OGUJF3TV			CC717411	OGUJF3TV		
246	15.4	85.6	584	2	BB667499	601442591			BB667499	601442591		
C 247	15.4	85.6	585	4	B1368953	RE56305.5			B1368953	RE56305.5		
248	15.4	85.6	585	7	CN482270	hw18C10.Y			CN482270	hw18C10.Y		
C 249	15.4	85.6	590	4	B1490893	603032132			B1490893	603032132		
C 250	15.4	85.6	591	4	B1233517	949011F12			B1233517	949011F12		
251	15.4	85.6	591	2	AW953223	ES7365293			AW953223	ES7365293		
252	15.4	85.6	591	7	CK004991	AGENCOURT			CK004991	AGENCOURT		
253	15.4	85.6	594	2	BF024553	PBP 418 L			BF024553	PBP 418 L		
254	15.4	85.6	595	2	BF024230	PVP 593 L			BF024230	PVP 593 L		
255	15.4	85.6	595	8	BF024386	PBP 145 L			BF024386	PBP 145 L		
256	15.4	85.6	596	2	BZ992625	PUGH25TD			BZ992625	PUGH25TD		
C 257	15.4	85.6	597	1	CG761852	CDMMB033			CG761852	CDMMB033		
C 258	15.4	85.6	599	1	AA264286	LD07923.5			AA264286	LD07923.5		
259	15.4	85.6	603	5	B0638142	hd19a02.Y			B0638142	hd19a02.Y		
C 260	15.4	85.6	604	1	AA246917	LD05956.5			AA246917	LD05956.5		
C 261	15.4	85.6	605	4	BM686015	UI-E-CRO-			BM686015	UI-E-CRO-		
C 262	15.4	85.6	606	1	A1456248	LD36135.5			A1456248	LD36135.5		
C 263	15.4	85.6	609	4	B1633351	SD27421.5			B1633351	SD27421.5		
C 264	15.4	85.6	611	2	BE706378	RC1-HW025			BE706378	RC1-HW025		
C 265	15.4	85.6	611	2	B1213388	RE18775.5			B1213388	RE18775.5		
266	15.4	85.6	611	5	BQ807784	NISC_KK10			BQ807784	NISC_KK10		
267	15.4	85.6	612	5	AM964616	INSF11.Y			AM964616	INSF11.Y		
268	15.4	85.6	613	2	BU786506	EST376689			BU786506	EST376689		
269	15.4	85.6	613	2	BE313201	BX335925			BE313201	BX335925		
C 270	15.4	85.6	613	5	BX335925	BX335925			BX335925	BX335925		
C 271	15.4	85.6	615	4	B1389794	949053H12			B1389794	949053H12		
C 272	15.4	85.6	617	1	BE256395	601117857			BE256395	601117857		
C 273	15.4	85.6	619	2	AA391558	LD10559.5			AA391558	LD10559.5		
C 274	15.4	85.6	620	4	B1170548	RE12053.5			B1170548	RE12053.5		
C 275	15.4	85.6	620	4	BM763152	K-EST0044			BM763152	K-EST0044		
276	15.4	85.6	621	7	CN742130	SAT-US002			CN742130	SAT-US002		
277	15.4	85.6	630	1	BM111668	EST559204			BM111668	EST559204		
C 278	15.4	85.6	630	4	AA803515	GM13370.5			AA803515	GM13370.5		
C 279	15.4	85.6	632	2	BF024292	PBP 008 L			BF024292	PBP 008 L		
C 280	15.4	85.6	638	2	BF346432	602020273			BF346432	602020273		
281	15.4	85.6	639	2	BF024108	PVP 399 L			BF024108	PVP 399 L		
C 282	15.4	85.6	640	1	A1295263	LP08845.5			A1295263	LP08845.5		
C 283	15.4	85.6	640	4	B1486287	RE69703.5			B1486287	RE69703.5		
C 284	15.4	85.6	650	5	BQ808015	NISC_KK12			BQ808015	NISC_KK12		
C 285	15.4	85.6	651	1	A1260669	LP04622.5			A1260669	LP04622.5		
286	15.4	85.6	659	4	B1196759	602755394			B1196759	602755394		
287	15.4	85.6	661	4	BM017318	603643982			BM017318	603643982		
288	15.4	85.6	662	7	CK717642	17526_Swo			CK717642	17526_Swo		
C 289	15.4	85.6	666	6	B1171911	RE13809.5			B1171911	RE13809.5		
C 290	15.4	85.6	666	6	CB111111	K-EST0152			CB111111	K-EST0152		
291	15.4	85.6	667	4	B1823953	603039135			B1823953	603039135		
292	15.4	85.6	667	9	CG181690	PUIHQ30TB			CG181690	PUIHQ30TB		
293	15.4	85.6	669	6	CB113008	K-EST0154			CB113008	K-EST0154		
294	15.4	85.6	670	4	BG773349	602722129			BG773349	602722129		
295	15.4	85.6	670	4	B1561480	603256167			B1561480	603256167		
C 296	15.4	85.6	672	7	W22059	63P5_Human			W22059	63P5_Human		
297	15.4	85.6	672	9	AG044078	Pan_txcg1			AG044078	Pan_txcg1		
298	15.4	85.6	673	2	BE907298	601500006			BE907298	601500006		
299	15.4	85.6	674	7	CN296314	170004708			CN296314	170004708		
300	15.4	85.6	675	4	BG700653	602682270			BG700653	602682270		
C 301	15.4	85.6	677	6	CA781350	015F06AF			CA781350	015F06AF		
C 302	15.4	85.6	680	4	BG820014	602782548			BG820014	602782548		
C 303	15.4	85.6	682	8	A2521293	RPCT-11-4			A2521293	RPCT-11-4		
C 304	15.4	85.6	682	7	CN402429	170005999			CN402429	170005999		
305	15.4	85.6	686	4	BG826767	602751224			BG826767	602751224		
306	15.4	85.6	690	4	B1596327	603243051			B1596327	603243051		
307	15.4	85.6	693	6	CB156629	K-EST0215			CB156629	K-EST0215		
308	15.4	85.6	696	7	CG392076	ZMMBCC057			CG392076	ZMMBCC057		
309	15.4	85.6	697	7	CK591840	esc_1_Yan			CK591840	esc_1_Yan		
C 310	15.4	85.6	698	4	B1372340	RE59280.5			B1372340	RE59280.5		
311	15.4	85.6	698	4	B1460475	603200973			B1460475	603200973		
312	15.4	85.6	698	6	CD871526	AZ02_118H			CD871526	AZ02_118H		
C 313	15.4	85.6	699	8	BZ021404	oe489d08			BZ021404	oe489d08		
C 314	15.4	85.6	700	8	AG158499	Pan_txcg1			AG158499	Pan_txcg1		
315	15.4	85.6	702	7	CN402420	170005830			CN402420	170005830		
316	15.4	85.6	705	4	B1759405	603043071			B1759405	603043071		
317	15.4	85.6	705	8	BZ021474	oe489d09			BZ021474	oe489d09		
318	15.4	85.6	706	4	BG704777	602688401			BG704777	602688401		
319	15.4	85.6	709	4	BG743670	602632842			BG743670	602632842		
320	15.4	85.6	710	4	B1596935	603242951			B1596935	603242951		
321	15.4	85.6	711	4	B1561401	603255168			B1561401	603255168		
322	15.4	85.6	718	1	AL557717	AL557717			AL557717	AL557717		
323	15.4	85.6	719	7	CK591868	esc_1_Yan			CK591868	esc_1_Yan		
324	15.4	85.6	725	4	BG703713	602686594			BG703713	602686594		
325	15.4	85.6	729	4	B1549063	603197041			B1549063	603197041		
326	15.4	85.6	734	4	BG723009	602699541			BG723009	602699541		
327	15.4	85.6	736	4	B1562580	603256513			B1562580	603256513		
328	15.4	85.6	737	5	BQ109492	imageqc_7			BQ109492	imageqc_7		
C 329	15.4	85.6	739	5	BQ870034	QGB7018.Y			BQ870034	QGB7018.Y		
330	15.4	85.6	740	4	BG570824	602591951			BG570824	602591951		
331	15.4	85.6	740	7	CN402427	170006008			CN402427	170006008		
332	15.4	85.6	741	4	BG753204	602732289			BG753204	602732289		
333	15.4	85.6	742	4	B1762198	603049102			B1762198	603049102		
334	15.4	85.6	748	4	BG327436	602426418			BG327436	602426418		
335	15.4	85.6	751	2	BE797043	601588087			BE797043	601588087		
336	15.4	85.6	752	4	BG700421	602680188			BG700421	602680188		
337	15.4	85.6	752	4	BM048451	603628131			BM048451	603628131		
338	15.4	85.6	753	9	CR230698	FOFWA3 B			CR230698	FOFWA3 B		
339	15.4	85.6	756	6	CF126302	UI-HF-ET0			CF126302	UI-HF-ET0		
340	15.4	85.6	758	4	BG471654	602513244			BG471654	602513244		
341	15.4	85.6	761	4	B1602722	603252068			B1602722	603252068		
342	15.4	85.6	763	4	BG913449	602811339			BG913449	602811339		
C 343	15.4	85.6	774	9	CG501062	OST44623			CG501062	OST44623		
344	15.4	85.6	780	2	BE545846	601071748			BE545846	601071748		
345	15.4	85.6	782	8	BZ255203	CH230-478			BZ255203	CH230-478		
346	15.4	85.6	786	5	BQ855467	QGB26116			BQ855467	QGB26116		
347	15.4	85.6	792	4	B1093371	602895751			B1093371	602895751		
348	15.4	85.6	797	7	CK706138	601089751			CK706138	601089751		
C 349	15.4	85.6	809	7	CK121149	204612_P1			CK121149	204612_P1		
C 350	15.4	85.6	811	4	B1916546	603178456			B1916546	603178456		
351	15.4	85.6	812	6	CD107220	AGENCOURT			CD107220	AGENCOURT		
C 352	15.4	85.6	816	8	BZ771256	mcb74D02			BZ771256	mcb74D02		
353	15.4	85.6	816	4	B1553513	603193294			B1553513	603193294		
354	15.4	85.6	817	2	BE904792	601498933						

390	15.4	85.6	869	1	AL536432	AL536432	463	15.4	85.6	1588	3	CR591088	CR591088 full-1eng
391	15.4	85.6	877	2	BE267279	BE267279	464	15.4	85.6	1592	3	CR599907	CR599907 full-1eng
392	15.4	85.6	877	2	BM047993	BM047993	465	15.4	85.6	1605	3	CR614623	CR614623 full-1eng
393	15.4	85.6	878	1	AL522006	AL522006	466	15.4	85.6	1626	3	CR603960	CR603960 full-1eng
394	15.4	85.6	880	6	BF247221	BF247221	467	15.4	85.6	1642	3	CR609654	CR609654 full-1eng
395	15.4	85.6	880	6	CD107479	CD107479	468	15.4	85.6	1653	3	CR625078	CR625078 full-1eng
396	15.4	85.6	881	5	BJ0725856	BJ0725856	469	15.4	85.6	1711	3	CR620589	CR620589 full-1eng
397	15.4	85.6	883	4	BI835449	BI835449	470	15.4	85.6	1716	3	CR617488	CR617488 full-1eng
398	15.4	85.6	885	4	BG537052	BG537052	471	15.4	85.6	1712	3	CR618858	CR618858 full-1eng
399	15.4	85.6	886	6	CD251096	CD251096	472	15.4	85.6	1712	3	CR618858	CR618858 full-1eng
400	15.4	85.6	889	5	CD106781	CD106781	473	15	83.3	208	1	AL3366846	AL3366846 full-1eng
401	15.4	85.6	890	5	BC518585	BC518585	474	15	83.3	273	9	CR344486	CR344486 full-1eng
402	15.4	85.6	890	5	BUB57368	BUB57368	475	15	83.3	404	2	AM127172	AM127172 full-1eng
403	15.4	85.6	892	4	BG675133	BG675133	476	15	83.3	420	1	AI629102	AI629102 full-1eng
404	15.4	85.6	897	6	CD386873	CD386873	477	15	83.3	420	8	AZ993298	AZ993298 full-1eng
405	15.4	85.6	897	7	CNS03787	CNS03787	478	15	83.3	421	8	AO167120	AO167120 full-1eng
406	15.4	85.6	899	4	BG113617	BG113617	479	15	83.3	462	4	BM694409	BM694409 full-1eng
407	15.4	85.6	899	4	BG623870	BG623870	480	15	83.3	475	1	AL369880	AL369880 full-1eng
408	15.4	85.6	900	5	BK324859	BK324859	481	15	83.3	498	1	AL923940	AL923940 full-1eng
409	15.4	85.6	902	5	BU180084	BU180084	482	15	83.3	500	8	AO224271	AO224271 full-1eng
410	15.4	85.6	903	5	BU904211	BU904211	483	15	83.3	514	2	AM559718	AM559718 full-1eng
411	15.4	85.6	903	5	BK847125	BK847125	484	15	83.3	553	9	CE267224	CE267224 full-1eng
412	15.4	85.6	904	8	BZ820113	BZ820113	485	15	83.3	555	6	CA589677	CA589677 full-1eng
413	15.4	85.6	905	2	BR024337	BR024337	486	15	83.3	593	7	CR283635	CR283635 full-1eng
414	15.4	85.6	909	5	BU339399	BU339399	487	15	83.3	609	8	BM727257	BM727257 full-1eng
415	15.4	85.6	912	4	BG428401	BG428401	488	15	83.3	621	9	CR487638	CR487638 full-1eng
416	15.4	85.6	914	2	BE619436	BE619436	489	15	83.3	624	9	CE236144	CE236144 full-1eng
417	15.4	85.6	915	5	BQ942387	BQ942387	490	15	83.3	653	9	AG137736	AG137736 full-1eng
418	15.4	85.6	917	2	BF529453	BF529453	491	15	83.3	670	6	CF068406	CF068406 full-1eng
419	15.4	85.6	920	5	BU542600	BU542600	492	15	83.3	672	9	AG067128	AG067128 full-1eng
420	15.4	85.6	932	2	BF347709	BF347709	493	15	83.3	680	9	CR487857	CR487857 full-1eng
421	15.4	85.6	932	2	BI602593	BI602593	494	15	83.3	680	9	CR487857	CR487857 full-1eng
422	15.4	85.6	937	6	CD513871	CD513871	495	15	83.3	702	2	AM129028	AM129028 full-1eng
423	15.4	85.6	941	2	BI0105602	BI0105602	496	15	83.3	718	2	AC954398	AC954398 full-1eng
424	15.4	85.6	944	2	BF024113	BF024113	497	15	83.3	702	2	AM129028	AM129028 full-1eng
425	15.4	85.6	945	2	BM472787	BM472787	498	15	83.3	718	2	AC954398	AC954398 full-1eng
426	15.4	85.6	947	8	BZ820107	BZ820107	499	15	83.3	750	9	CC566017	CC566017 full-1eng
427	15.4	85.6	955	2	BE741394	BE741394	500	15	83.3	751	7	CO132211	CO132211 full-1eng
428	15.4	85.6	957	2	BG164320	BG164320	501	15	83.3	763	7	CO132211	CO132211 full-1eng
429	15.4	85.6	966	5	BG115287	BG115287	502	15	83.3	778	4	BI857736	BI857736 full-1eng
430	15.4	85.6	968	5	BK419451	BK419451	503	15	83.3	778	4	BI857736	BI857736 full-1eng
431	15.4	85.6	969	4	BG696276	BG696276	504	15	83.3	778	4	BI857736	BI857736 full-1eng
432	15.4	85.6	971	3	CNS0A1JM	CNS0A1JM	505	15	83.3	778	4	BI857736	BI857736 full-1eng
433	15.4	85.6	971	7	CO579861	CO579861	506	15	83.3	778	4	BI857736	BI857736 full-1eng
434	15.4	85.6	983	1	AL539546	AL539546	507	15	83.3	778	4	BI857736	BI857736 full-1eng
435	15.4	85.6	984	1	AL530745	AL530745	508	15	83.3	778	4	BI857736	BI857736 full-1eng
436	15.4	85.6	989	9	CNS030LS	CNS030LS	509	15	83.3	778	4	BI857736	BI857736 full-1eng
437	15.4	85.6	991	5	BK409882	BK409882	510	15	83.3	778	4	BI857736	BI857736 full-1eng
438	15.4	85.6	995	5	BK398758	BK398758	511	15	83.3	778	4	BI857736	BI857736 full-1eng
439	15.4	85.6	1009	5	BQ880780	BQ880780	512	15	83.3	778	4	BI857736	BI857736 full-1eng
440	15.4	85.6	1012	4	BM562959	BM562959	513	15	83.3	778	4	BI857736	BI857736 full-1eng
441	15.4	85.6	1017	4	BI458505	BI458505	514	15	83.3	778	4	BI857736	BI857736 full-1eng
442	15.4	85.6	1024	5	BQ068489	BQ068489	515	15	83.3	778	4	BI857736	BI857736 full-1eng
443	15.4	85.6	1031	6	CD512978	CD512978	516	15	83.3	778	4	BI857736	BI857736 full-1eng
444	15.4	85.6	1052	4	BM921331	BM921331	517	15	83.3	778	4	BI857736	BI857736 full-1eng
445	15.4	85.6	1053	5	BQ067722	BQ067722	518	15	83.3	778	4	BI857736	BI857736 full-1eng
446	15.4	85.6	1060	5	EX84498	EX84498	519	15	83.3	778	4	BI857736	BI857736 full-1eng
447	15.4	85.6	1062	4	BG164254	BG164254	520	15	83.3	778	4	BI857736	BI857736 full-1eng
448	15.4	85.6	1064	1	AL541105	AL541105	521	15	83.3	778	4	BI857736	BI857736 full-1eng
449	15.4	85.6	1072	7	CK029809	CK029809	522	15	83.3	778	4	BI857736	BI857736 full-1eng
450	15.4	85.6	1075	6	CF272425	CF272425	523	15	83.3	778	4	BI857736	BI857736 full-1eng
451	15.4	85.6	1091	8	CC274773	CC274773	524	15	83.3	778	4	BI857736	BI857736 full-1eng
452	15.4	85.6	1156	4	BM473378	BM473378	525	15	83.3	778	4	BI857736	BI857736 full-1eng
453	15.4	85.6	1269	3	CNS0A0Y4	CNS0A0Y4	526	15	83.3	778	4	BI857736	BI857736 full-1eng
454	15.4	85.6	1345	3	CR590725	CR590725	527	15	83.3	778	4	BI857736	BI857736 full-1eng
455	15.4	85.6	1375	3	CNS09YTN	CNS09YTN	528	15	83.3	778	4	BI857736	BI857736 full-1eng
456	15.4	85.6	1378	8	CC261389	CC261389	529	15	83.3	778	4	BI857736	BI857736 full-1eng
457	15.4	85.6	1383	3	CNS0A0R1	CNS0A0R1	530	15	83.3	778	4	BI857736	BI857736 full-1eng
458	15.4	85.6	1386	3	CR621335	CR621335	531	15	83.3	778	4	BI857736	BI857736 full-1eng
459	15.4	85.6	1440	3	CR602177	CR602177	532	15	83.3	778	4	BI857736	BI857736 full-1eng
460	15.4	85.6	1516	3	CR609599	CR609599	533	15	83.3	778	4	BI857736	BI857736 full-1eng
461	15.4	85.6	1561	3	CR613569	CR613569	534	15	83.3	778	4	BI857736	BI857736 full-1eng
462	15.4	85.6	1564	3	CR601079	CR601079	535	15	83.3	778	4	BI857736	BI857736 full-1eng

## ALIGNMENTS

BB571570 237 bp mRNA linear EST 30-NOV-2000  
 BB571570 RIKEN full-length enriched, 10 day neonate skin Mus  
 musculus cDNA clone 4733402E01 5', mRNA sequence.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1-7-22 Suenito-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

**FEATURES**  
**SOURCE**

**Source**

**ORIGIN**

**IN**

二

二

**0000**

**0000**

CE

CE

## REFERENCE

## REFERENCE

## COMMENT

## COMMENT

Seq pr:

## Seq pr:

the mis

the mis

9

9

11

11

## ORGANIS

## ORGANIS

## AUTHORS

## AUTHORS

JOURNAL

## JOURNAL



Db 108 CGACTGTGAATCTTCAT 91

RESULT 6  
LOCUS BE127536/c 407 bp mRNA linear EST 15-JUN-2000  
DEFINITION DEB1285 Rat Lambda ZAP Express Library Rattus norvegicus cDNA 5',  
mRNA sequence.  
ACCESSION BE127536  
VERSION BE127536.1 GI:8550273  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS Sleeman, M.A., Morrison, J.G., Strachan, L., Kumble, K.D., Glenn, M.P.,  
McGrath, A., Grierison, A., Havukkala, I., Tan, P.L.J. and Watson, J.D.  
TITLE Expressed sequence tags of cDNA clones from rat dermal papilla  
cells  
JOURNAL Unpublished (2000)  
COMMENT Contact: Sleeman MA  
Biology  
Genesis Research and Development Corporation Limited  
P.O. Box 50, 1 Fox St, Parnell, Auckland, New Zealand  
Tel: 0064 9 373 5600  
Fax: 0064 9 373 2189  
Email: m.sleeman@genesis.co.nz  
Seq primer: T3 forward  
High quality sequence stop: 407.  
Location/Qualifiers  
1..407  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Dark-Agouti"  
/db\_xref="taxon:10116"  
/tissue\_type="vibrissae"  
/cell\_type="dermal papilla"  
/clone\_lib="Rat Lambda ZAP Express Library"

ORIGIN  
Query Match 100.0%; Score 18; DB 2; Length 407;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTTCAT 18  
Db 121 CGACTGTGAATCTTCAT 104

RESULT 7  
LOCUS CB795162/c 427 bp mRNA linear EST 16-MAY-2003  
DEFINITION AMGNMNC:SRPG2-00011-C10-A srpg2 (10238) Rattus norvegicus cDNA  
clone srpg2-00011-cl0 5', mRNA sequence.  
ACCESSION CB795162  
VERSION CB795162.1 GI:29883639  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS Angen, Inc  
TITLE Angen EST Program  
JOURNAL Angen Rat EST Program  
COMMENT Contact: Dan Fitzpatrick  
Angen, Inc  
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00011 row: c column: 10.

FEATURES  
source Location/Qualifiers  
1..427  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone\_lib="srpg2-00011-cl0"  
/tissue\_type="pancreal gland brain"  
/clone\_lib="srpg2 (10238)"  
/note="Vector: pSPORT1, Site\_1: SalI, Site\_2: NotI, pancreal  
gland brain region"

ORIGIN  
Query Match 100.0%; Score 18; DB 6; Length 427;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTTCAT 18  
Db 121 CGACTGTGAATCTTCAT 104

RESULT 8  
LOCUS CD372427/c 429 bp mRNA linear EST 29-MAY-2003  
DEFINITION UI-R-G00-csf-b-03-0-UI.r1 UI-R-G00 Rattus norvegicus cDNA clone  
UI-R-G00-csf-b-03-0-UI 5', mRNA sequence.  
ACCESSION CD372427  
VERSION CD372427.1 GI:31156517  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: James Lin, University of Iowa  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/rat.html>  
Seq primer: M13 REVERSE.

FEATURES  
source Location/Qualifiers  
1..429  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone\_lib="UI-R-G00-csf-b-03-0-UI"  
/tissue\_type="Whole embryo"  
/dev\_stage="embryo 13dpc"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-R-G00"  
/note="Vector: pYX-AseI, Site 1: EcoR I; Site 2: Not I;  
UI-R-G00 is a cDNA library containing the following  
tissue(s): rat whole embryo 13dpc. The library was  
constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. Denatured RNA was size  
fractionated on a 1% agarose gel. First strand cDNA  
synthesis was primed with oligo-dT primer containing a Not

I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pTX-Lac vector. The library tag sequence located between the Not I site and the polyA tail is CAGCTCTACT. This library was created for the University of Iowa Program for Rat Gene Discovery and Mapping (Val Sheffield, Bento Soares and Tom Casavant)."

## ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 429;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGATCCTCCAT 18  
|||||  
Db 239 CGACTGTGATCCTCCAT 222

RESULT 9  
CB732629 471 bp mRNA linear EST 11-APR-2003  
LOCUS

DEFINITION AMNCNC:MBB3-00106-H11-A rat brain E15 (10374) Rattus norvegicus  
CB732629  
CB732629.1 GI:29799792

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus (Norway rat)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Amgen, Inc  
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00106 row: h column: 11.  
Location/Qualifiers  
1. 471  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="mbb3-00106-h11"  
/tissue\_type="brain E15"  
/clone\_lib="rat brain E15 (10374)"  
/note="Vector: pECB; Site\_1: BstXI; Site\_2: NotI; rat brain E15"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Amgen EST Program  
Amgen Rat EST Program  
Unpublished (2003)  
Contact: Dan Fitzpatrick

## FEATURES

source  
1. 471  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="mbb3-00106-h11"  
/tissue\_type="brain E15"  
/clone\_lib="rat brain E15 (10374)"  
/note="Vector: pECB; Site\_1: BstXI; Site\_2: NotI; rat brain E15"

## ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 471;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGATCCTCCAT 18  
|||||  
Db 200 CGACTGTGATCCTCCAT 183

RESULT 10  
BE128412 479 bp mRNA linear EST 15-JUN-2000  
LOCUS

DEFINITION DEPA2162 Rat lambda ZAP Express Library Rattus norvegicus cDNA 5',  
mRNA sequence.  
BE128412  
BE128412.1 GI:8551165

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE  
1 (bases 1 to 479)

## AUTHORS

Sleeman, M.A., Morrison, J.G., Strachan, L., Kumble, K.D., Glenn, M.P.,  
McCrack, A., Grierson, A., Havukkala, I., Tan, P.L.J. and Watson, J.D.  
Expressed sequence tags of cDNA clones from rat dermal papilla  
cells  
Unpublished (2000)

JOURNAL  
COMMENT  
Contact: Sleeman MA

## BIOLOGY

Genesis Research and Development Corporation Limited  
P.O. Box 50, 1 Fox St, Parnell, Auckland, New Zealand  
Tel: 0064 9 373 5600  
Fax: 0064 9 373 2189  
Email: m.sleeman@genesis.co.nz

Seq primer: T3 forward  
High quality sequence stop: 479.  
Location/Qualifiers  
1. 479  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Dark-Agouti"  
/db\_xref="taxon:10116"  
/tissue\_type="vibrissae"  
/cell\_type="dermal papilla"  
/clone\_lib="Rat Lambda ZAP Express Library"

## FEATURES

## source

## ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 479;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGATCCTCCAT 18  
|||||  
Db 121 CGACTGTGATCCTCCAT 104

RESULT 11  
BF554812 559 bp mRNA linear EST 12-DEC-2000  
LOCUS

DEFINITION BF554812  
BF554812.1 GI:11664584  
BF554812  
BF554812.1 GI:11664584

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus (Norway rat)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
CDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Clones will be available through Research Genetics (www.resgen.com)  
This clone is also available through the I.M.A.G.E. Consortium at  
HLNI (info@image.llnl.gov). IMAGE ID= 1770040  
Seq primer: M3 forward.  
Location/Qualifiers  
1. 559  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"

Normalization and subtraction: two approaches to facilitate gene

discovery  
Genome Res. 6 (9), 791-806 (1996)

## COMMENT



/strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-E0-cb-e-01-0-UI"  
 /dev\_stage="embryonic"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_1lb="UI-R-E0"  
 /note="Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; This library consists of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture."

## ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18  
 Db 127 CGACTGGAATCCTCCAT 110

RESULT 12  
 CB609941/c 559 bp mRNA linear EST 16-MAY-2003  
 LOCUS AMGNNUC:URTG1-00001-E7-A urtgl (13981) Rattus norvegicus cDNA clone  
 DEFINITION urtgl-00001-e7 5', mRNA sequence.

ACCESSION CB609941  
 VERSION CB609941.1 GI:29549590  
 KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 559)

AUTHORS Angen EST Program.

TITLE Angen Rat EST Program

JOURNAL Unpublished (2003)

COMMENT Contact: Dan Fitzpatrick

Angen, Inc  
 One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00001 row: e column: 7.

Location/Qualifiers

1..559  
 /organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="urtgl-00001-e7"

/clone\_1lb="urtgl (13981)"

/note="Vector: pSPORT1; Rat toxicology library Rearranged  
 from internal pSPORT vector"

## ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18  
 Db 113 CGACTGGAATCCTCCAT 96

## RESULT 13

CF978067/c

LOCUS CF978067 562 bp mRNA linear EST 24-JUN-2004  
 DEFINITION FA11354.R Rat retinal ganglion cell Rattus norvegicus cDNA, mRNA  
 sequence.

ACCESSION CF978067

VERSION CF978067.1 GI:49173525  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 562)

AUTHORS Parkes,R.H., Qian,J., Goldberg,J.L., Quigley,H.A. and Zack,D.J.

TITLE Gene Expression Profiling of Highly Purified Rat Retinal Ganglion

CELLS Cells

JOURNAL Unpublished (2003)

COMMENT Contact: Parkes RH

Department of Ophthalmology

Johns Hopkins University School of Medicine

600 North Wolfe Street, Baltimore, MD 21287, USA

Tel: 410 502 5230

Fax: 410 502 5382

Email: rfarkes@jhmi.edu.

## FEATURES

Location/Qualifiers

1..562  
 /organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/tissue\_type="Retinal Ganglion Cells"

/lab\_host="DH10B"

/clone\_1lb="Rat retinal ganglion cell"

/note="Organ: Eye; Vector: pDNR-LIB; Site\_1: SfiI; Site\_2:

SfiI; The library was constructed from purified rat

retinal ganglion cells. The Creator SMART cDNA library

method (Clontech) was used. EST analysis was performed on

the unamplified, non-normalized, non-subtracted library."

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 87;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18  
 Db 132 CGACTGGAATCCTCCAT 115

## RESULT 14

CB583399/c

LOCUS CB583399 589 bp mRNA linear EST 03-APR-2003  
 DEFINITION AMGNNUC:NRDGI-00174-D10-A nrdgl (10855) Rattus norvegicus cDNA  
 clone nrdgl-00174-d10 5', mRNA sequence.

ACCESSION CB583399

VERSION CB583399.1 GI:29528865

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 589)

AUTHORS Angen EST Program.

TITLE Angen Rat EST Program

JOURNAL Unpublished (2003)

COMMENT Contact: Dan Fitzpatrick

Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00174 row: d column: 10.

Location/Qualifiers

1..589  
 /organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="nrdgl-00174-d10"

/tissue\_type="Dorsal Root Ganglia"

ORIGIN /clone\_1lb="nrldg (10855)"  
/note="Vector: pSPORT1; Site\_1: Sal1; Site\_2: NotI; rat  
dorsal root ganglia"

Query Match 100.0%; Score 18; DB 6; Length 589;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCAT 18  
Db 103 CGACTGTGAATCCTCCAT 86

RESULT 15  
CO400616/c 715 bp mRNA linear EST 01-JUL-2004  
LOCUS AGENCOURT 27533120 NIH\_MGC\_254 Rattus norvegicus cDNA clone  
IMAGE:7312060 5', mRNA sequence.  
ACCESSION CO400616  
VERSION CO400616  
KEYWORDS EST  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 715)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cga@db-remail.nih.gov  
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
College of Wisconsin  
CDNA Library Preparation: Express Genomics  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: L1AM15360 row: 0 column: 02  
High quality sequence stop: 655.  
Location/Qualifiers

## FEATURES

1..715  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="IMAGE:7312060"  
/sex="both"  
/tissue\_type="Brain - Pooled from several tissues from one  
or more individuals"  
/lab\_host="DH10B TONa"  
/clone\_1lb="NIH\_MGC\_254"  
/note="Organ: brain/CNS; Vector: pExpress-1; Site 1:  
ECORV; Site 2: NotI; RNA obtained from brain tissue of 8  
wk old animal. Tissues were snap-frozen and kept at -80C  
before RNA extraction and purification (TRI-reagent  
method). cDNA was primed using oligo-dT primer:  
5'-pgACTAGTTCGATCGGAGCGCGCC(T)25-3' and cloned into  
the ECORV/NotI sites of pExpress-1. Size-selection >1.25kb  
resulted in an average insert size of 2.18 kb. This  
primary library is not normalized (normalized library is  
NIH\_MGC\_255) and was constructed by Express Genomics  
(Frederick, MD). Note: this is a NIH\_MGC library"

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 715;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 155 CGACTGTGAATCCTCCAT 138

RESULT 16  
CO394200/c 744 bp mRNA linear EST 01-JUL-2004  
LOCUS AGENCOURT 27533383 NIH\_MGC\_254 Rattus norvegicus cDNA clone  
IMAGE:7311837 5', mRNA sequence.  
ACCESSION CO394200  
VERSION CO394200  
KEYWORDS EST  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cga@db-remail.nih.gov  
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
College of Wisconsin  
CDNA Library Preparation: Express Genomics  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: L1AM15360 row: e column: 19  
High quality sequence stop: 666.  
Location/Qualifiers

## FEATURES

1..744  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="IMAGE:7311837"  
/sex="both"  
/tissue\_type="Brain - Pooled from several tissues from one  
or more individuals"  
/lab\_host="DH10B TONa"  
/clone\_1lb="NIH\_MGC\_254"  
/note="Organ: brain/CNS; Vector: pExpress-1; Site 1:  
ECORV; Site 2: NotI; RNA obtained from brain tissue of 8  
wk old animal. Tissues were snap-frozen and kept at -80C  
before RNA extraction and purification (TRI-reagent  
method). cDNA was primed using oligo-dT primer:  
5'-pgACTAGTTCGATCGGAGCGCGCC(T)25-3' and cloned into  
the ECORV/NotI sites of pExpress-1. Size-selection >1.25kb  
resulted in an average insert size of 2.18 kb. This  
primary library is not normalized (normalized library is  
NIH\_MGC\_255) and was constructed by Express Genomics  
(Frederick, MD). Note: this is a NIH\_MGC library"

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 744;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCAT 18  
Db 155 CGACTGTGAATCCTCCAT 138

RESULT 17  
CK482053/c

```

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7108804"
/tissue_type="kidney, pooled"
/lab_host="DH10B Tona"
/clone_1fb="NIH MGC 235"
/notes="Organ: kidney; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; RNA obtained from pooled kidney tissue from a mix of male and female animals at 8 wk old. Tissues were snap-frozen before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer: 5'-gacatgatttgaatccgacgacgccc(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb resulted in an average insert size of 2.2 kb. This primary library is non-normalized (normalized primary library is NIH MGC 236) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH MGC library."

```

REFERENCE	Rattus.			
AUTHORS	1 (bases 1 to 797)			
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: <a href="mailto:cgaps-r@mail.nih.gov">cgaps-r@mail.nih.gov</a> Tissue Procurement: Howard Jacobs cDNA Library Preparation: Express Genomics cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LLM15013 row: j column: 13 High quality sequence set: 718. Location/Qualifiers			
FEATURES	1. .797 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" /clone="IMAGE:7124271" /tissue_type="lung, pooled" /lab_host="DH10B Tona" /clone_id="NIH MGC 232" /note="Organ: lung; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI. RNA obtained from pooled lung tissue from a mix of male and female animals at 8 wk old. Tissues were snap-frozen and kept at -80C for two days before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer: 5'-pGACTGTCGTGATGTCGAGCGCGCCGCT(25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb resulted in an average insert size of 2.3 kb. This primary library is normalized (non-normalized primary library is NIH MGC 231) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library."			
ORIGIN				
Query Match	100.0%;	Score 18;	DB 7;	Length 797;
Best Local Similarity	100.0%;	Pred. No. 91;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Or	1 CGACTGTGATTCCTCCAT 18       176 CGACTGTGATTCCTCCAT 159			
Db	176 CGACTGTGATTCCTCCAT 159			
RESULT 19				
COS60520/c	COS60520 802 bp mRNA 11near EST 19-JUL-2004			
LOCUS	AGENCOURT_28621896 NIH_MGC_250 Rattus norvegicus cDNA clone			
DEFINITION	IMAGE:7384467 5', mRNA sequence.			
ACCESSION	COS60520			
VERSION	COS60520.1 GI:50373116			
KEYWORDS	EST.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE	1 (bases 1 to 802)			
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: <a href="mailto:cgaps-r@mail.nih.gov">cgaps-r@mail.nih.gov</a>			

Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical College of Wisconsin  
 cDNA Library Preparation: Open Biosystems  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 Clone Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LHAM1549 row: h column: 01  
 High quality sequence start: 24  
 High quality sequence stop: 725.  
 Location/Qualifiers  
 1..802

## FEATURES

source

/organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7384467"  
 /lab\_host="DH10B TONa"  
 /clone\_1lb="NIH MGC 250"  
 /note="Organ: thymus; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; RNA obtained from testis tissue of 8 wk old animal. Tissues were snap-frozen and kept at -80C before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer:  
 5'-pGACTAGTCTAGATCGAGCGCGCCGCT(7)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 1.9 kb. This is a primary library (normalized library is NIH MGC 251) and was constructed by Open Biosystems. Note: this is a NIH MGC library"

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 802;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTTCAT 18  
 |||||  
 Db 150 CGACTGTGAATCTTCAT 133

RESULT 20  
 CB315809/c 845 bp mRNA linear EST 04-MAR-2003  
 LOCUS AGNCCOURT 11525266 NICHD Rr Plt1 Rattus norvegicus cDNA clone  
 DEFINITION IMAGE:6888160 5', mRNA sequence.  
 CB315809  
 ACCESSION CB315809.1 GI:28840044  
 VERSION  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 845)  
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: John C. Marshall, M.D., Ph.D  
 cDNA Library Preparation: CLONTECH  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LHCW3143 row: p column: 15  
 High quality sequence stop: 390.  
 Location/Qualifiers  
 1..845  
 /organism="Rattus norvegicus"

## FEATURES

source

/organism="Rattus norvegicus"

/mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:6888160"  
 /tissue\_type="Pituitary"  
 /lab\_host="DH10B"  
 /clone\_1lb="NICHD Rr Plt1"  
 /note="Vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI; 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.23 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech laboratories (Palo Alto, CA)."

## ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 845;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTTCAT 18  
 |||||  
 Db 188 CGACTGTGAATCTTCAT 171

RESULT 21  
 CK602723/c 846 bp mRNA linear EST 22-JAN-2004  
 LOCUS AGNCCOURT 17901070 NIH MGC 234 Rattus norvegicus cDNA clone  
 DEFINITION IMAGE:7193583 5', mRNA sequence.  
 CK602723  
 ACCESSION CK602723.1 GI:41116034  
 VERSION  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 846)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics / NIH  
 National Cancer Institute  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Howard Jacobs  
 cDNA Library Preparation: Express Genomics  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LHAM15056 row: b column: 13  
 High quality sequence stop: 681.  
 Location/Qualifiers  
 1..846  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7193583"  
 /tissue\_type="heart, pooled"  
 /lab\_host="DH10B TONa"  
 /clone\_1lb="NIH MGC 234"  
 /note="Organ: heart; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; RNA obtained from pooled heart tissue from a mix of male and female animals at 8 wk old. Tissues were snap-frozen and kept at -80C for two days before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer:  
 5'-pGACTAGTCTAGATCGAGCGCGCCGCT(7)25-3' and cloned into

## FEATURES

source

the ECoRV/NotI sites of pExpress-1. Size-selection >1.4kb resulted in an average insert size of 2.2 kb. This primary library is normalized (non-normalized primary library is NIH\_MGC\_233) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 846;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCTCCAT 18  
|||||  
77 CGACTGTGATCTCTCCAT 60

Db

RESULT 22  
BU896918 537 bp mRNA linear EST 17-OCT-2002  
LOCUS X047G05 Populus wood cDNA library Populus tremula x Populus tremuloides cDNA 5 prime, mRNA sequence.

ACCESSION BU896918 GI:24108125  
VERSION BU896918.1  
KEYWORDS Populus tremula x Populus tremuloides  
SOURCE Populus tremula x Populus tremuloides  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 537)  
AUTHORS Umeberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.  
TITLE The poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries  
JOURNAL Unpublished (2002)  
COMMENT Contact: BHALERAO RUPALI R.  
Umea Plant Science Center  
Department of Plant Physiology  
University of Umea, 901 87 Umea, Sweden  
Tel: +46 90 786 5279  
Fax: +46 90 786 6676  
Email: rupali.bhalerao@plantphys.umu.se.  
Location/Qualifiers

FEATURES  
SOURCE 1..537  
/organism="Populus tremula x Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:47664"  
/tissue\_type="wood"  
/clone\_lib="Populus wood cDNA library"

## ORIGIN

Query Match 94.4%; Score 17; DB 5; Length 537;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCTCA 17  
|||||  
434 CGACTGTGATCTCTCA 418

Db

RESULT 23  
CD829200/c 717 bp mRNA linear EST 10-JUL-2003  
LOCUS BM40.041G04F011207 BM40 Brassica napus cDNA clone BM40041G04, mRNA sequence.

ACCESSION CD829200 GI:32511140  
VERSION CD829200.1  
KEYWORDS EST.  
SOURCE Brassica napus (rape)  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 717)

AUTHORS Genoplante.  
TITLE Genoplante, a major partnership french program in plant genomics  
JOURNAL Unpublished (2003)  
COMMENT Contact: Genoplante  
Genoplante  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

## FEATURES

## SOURCE

1..717  
/organism="Brassica napus"  
/mol\_type="mRNA"  
/cultivar="Jet Neuf"  
/db\_xref="taxon:3708"  
/clone="BM40041G04"  
/tissue\_type="seed"  
/clone\_lib="BM40"

## ORIGIN

Query Match 94.4%; Score 17; DB 6; Length 717;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGATCTCTCCAT 18  
|||||  
650 GACTGTGATCTCTCCAT 634

Db

RESULT 24  
BZ186569/c 726 bp DNA linear GSS 11-OCT-2002  
LOCUS CH230-340J6.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone  
DEFINITION CH230-340J6, genomic survey sequence.  
ACCESSION BZ186569  
VERSION BZ186569.1 GI:23838462  
KEYWORDS GSS.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 726)  
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsagaye, G., Geer, K., Shvartsbeyn, A., Gebregorgis, B., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P., and Fraser, C.M.  
TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment Unpublished (1999)  
JOURNAL Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@igf.org

Clones are derived from the rat BAC library CHORI-230 (<http://www.chori.org/bacpac/rac230.htm>). For BAC library availability, please contact Pieter de Jong ([pdejong@mail.cho.org](mailto:pdejong@mail.cho.org)). Clones may be purchased from BACPAC Resources ([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end page: [http://www.tigr.org/cdb/bac\\_ends/rac/bac\\_end\\_intro.html](http://www.tigr.org/cdb/bac_ends/rac/bac_end_intro.html)  
Plate: 340 row: J column: 6  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers

## FEATURES

## SOURCE

1..726  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SNHed/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-340J6"

/sex="Female"  
/cell\_type="Brain"  
/clone\_lib="CHORI-230 Segment 2"  
/note="Vector: pFARBA1.3; Site 1: Mbol; Site 2: Mbol;  
CHORI-230 Rat (BN/SmHsd/MCM) BAC library produced by  
Pleter de Jong"

## ORIGIN

Query Match 94.4%; Score 17; DB 8; Length 726;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
|||||  
305 GACTGTGAATCCTCCAT 289

## Db

RESULT 25  
LOCUS CNS0591U 1075 bp DNA linear GSS 01-SEP-2000  
DEFINITION Tetradodon nigroviridis genome survey sequence T7 end of clone  
015B23 of library B from Tetradodon nigroviridis, genomic survey  
sequence.

ACCESSION AL326667.1 GI:8220256  
VERSION AL326667.1  
KEYWORDS GSS; genome survey sequence.  
SOURCE Tetradodon nigroviridis  
ORGANISM Tetradodon nigroviridis  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorphi; Acanthopterygii; Percormorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae; Tetraodon.

## REFERENCE

1 Roest Crolius, H., Jailton, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Bernot, A., Fitzames, C., Wincker, P., Brothier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.  
Estimate of human gene number provided by genome-wide analysis  
using Tetradodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20296633  
10835645  
2  
Roest Crolius, H., Jailton, O., Dasilva, C., Ozouf-Costaz, C.,  
Fitzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
Saurin, W., Bernot, A. and Weissenbach, J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetradodon nigroviridis  
Genome Res. 10 (7), 939-949 (2000)

## REFERENCE

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
10899143  
3 (bases 1 to 1075)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
Web : www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetradodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetradodon.

## REFERENCE

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
10899143  
3 (bases 1 to 1075)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
Web : www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetradodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetradodon.

## REFERENCE

## COMMENT

BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
Web : www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetradodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetradodon.

## COMMENT

location/Qualifiers  
1. 1075  
/organism="Tetradodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="015B23"  
/clone\_lib="B"  
/note="Genoscope sequence ID : C0AB015CA12C1-end : T7"

## FEATURES

## source

## ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 1075;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
|||||  
198 GACTGTGAATCCTCCAT 214

## Db

RESULT 26  
LOCUS CK334427/c 175 bp mRNA linear EST 22-DEC-2003  
DEFINITION H3026D07-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
H3026D07 5', mRNA sequence.

ACCESSION CK334427  
VERSION CK334427.1 GI:40290040  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS Tanaka, T.S., Jaradat, S.A., Lam, M.K., Kargul, G.J., Wang, X.,  
1 (bases 1 to 175)  
Wood, W.H., III, Becker, K.G. and Ko, M.S.H.  
Genome-wide expression profiling of mid-gestation placenta and  
embryo using a 15,000 mouse developmental cDNA microarray  
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

## AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20381348  
10922068  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@nigam.nih.gov  
Plate: H3026 row: D column: 07  
Seq primer: M13 Reverse  
High quality sequence stop: 175  
POLYA=No.

## COMMENT

location/Qualifiers  
1. 175  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taeST:H3026D07-5"  
/db\_xref="taxon:10090"  
/clone="H3026D07"  
/sex="Clones arrayed from a variety of cDNA libraries"  
/dev stage="Clones arrayed from a variety of cDNA libraries"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse 15K cDNA Clone Set"  
/note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; This  
clone is among a rearranged set of 15,247 clones from 11  
embryo cDNA libraries (including preimplantation stage  
embryos from unfertilized egg to blastocyst, embryonic  
part of E7.5 embryos, extraembryonic part of E7.5  
newborn ovary cDNA library. Average insert size 1.5 kb.  
All source libraries are cloned unidirectionally with  
Oligo(dT)-Not primers. References include: (1)  
Genome-wide expression profiling of mid-gestation  
placenta and embryo using a 15,000 mouse developmental  
cDNA microarray, 2000. Proc. Natl. Acad. Sci. U.S.A. 97:  
9127-9132; (2) Large-scale cDNA analysis reveals phased  
gene expression patterns during preimplantation mouse  
development, 2000. Development, 127: 1737-1749; (3)  
Genome-wide mapping of unselected transcripts from  
extraembryonic tissue of 7.5-day mouse embryos reveals  
enrichment in the t-complex and under-representation on  
the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

## FEATURES

## source

Query Match 91.1%; Score 16.4; DB 7; Length 175;  
Best Local Similarity 94.4%; Pred. No. 5.6e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## ORIGIN



/tissue\_type="embryonic body between diaphragm region and neck"  
 /dev\_stage="12 days embryo"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck"  
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCTCGACGCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGACGCTTTTCTTTTNNATTCCTCCCTCC 3']. cDNA was cleaved with XhoI and BamHI."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 233;  
 Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGATCTCTCCAT 18  
 |||||  
 Db 136 CGACTGATCTCTCCAT 119

RESULT 29  
 BB604900/c 245 bp mRNA linear EST 05-DEC-2000  
 LOCUS BB604900 RIKEN full-length enriched, 0 day neonate lung Mus  
 DEFINITION musculus cDNA clone E030007G03 5', mRNA sequence.  
 ACCESSION BB604900  
 VERSION BB604900.1 GI:11556302  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 245)

Alizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Hanagaki, T., Hayatsu, N., Hirooka, T., Hirozane, T., Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Iwata, M., Kawai, J., Kojima, Y., Kono, H., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T., Watabiki, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Yamamatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Alizawa, K. et al. 2000)  
 Unpublished (2000)

## JOURNAL

## COMMENT

LABORATORY for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermoregulation and thermoadaptation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Iwata, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

## FEATURES

## SOURCE

Location/Qualifiers  
 1. 245

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="E030007G03"  
 /tissue\_type="lung"  
 /dev\_stage="0 day neonate"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, 0 day neonate lung"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGACGCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGACGCTTTTCTTTTNNATTCCTCCCTCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLX I."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 245;  
 Best Local Similarity 94.4%; Pred. No. 5.9e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGATCTCTCCAT 18  
 |||||  
 Db 141 CGACTGATCTCTCCAT 124

RESULT 30  
 CN674653/c 252 bp mRNA linear EST 17-MAY-2004  
 LOCUS CN674653 A0951H02-5 NIA Mouse Embryonic Stem (ES) cell (11f+, 48 h, high density) cDNA library (Long) Mus musculus cDNA clone NIA:A0951H02  
 DEFINITION IMAGE:3077173 5', mRNA sequence.  
 ACCESSION CN674653  
 VERSION CN674653.1 GI:47441104  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 252)

Sharov, A.A., Piao, Y., Matoba, R., Dudekula, D.B., Qian, Y., VanBuren, V., Falco, G., Martin, P.R., Stagg, C.A., Baasey, U.C., Wang, Y., Carter, M.G., Hamatani, T., Alba, K., Akutsu, H., Sharova, L., Tanaka, T.S., Kimber, W.L., Yoshikawa, T., Jaradat, S.A., Pantano, S., Nagaraia, R., Boheler, K.R., Taub, D., Hodges, R.J., Longo, D.L., Schlessinger, D., Keller, J., Klotz, E., Kelsae, G., Umezawa, A., Vescovi, A.L., Rossant, J., Kumach, T., Hogan, B.L., Curci, A., D'Uso, M., Kelsae, J., Hide, W. and Ko, M.S.  
 Transcriptional analysis of mouse stem cells and early embryos  
 Proc Biol. 1 (3), 410-419 (2003)  
 Contact: Dawood B. Dudekula

## JOURNAL

## COMMENT

LABORATORY of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov



Plate: A0951 row: H column: 02  
 Seq primer: M13 Reverse  
 High quality sequence stop: 252  
 POLYA-No.

# FEATURES

source  
 1. 252  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129Sv/EvTac"  
 /db\_xref="nlaEST:A0951H02-5"  
 /db\_xref="taxon:10090"  
 /clone="NIA:A0951H02 IMAGE:3077173"  
 /sex="Male"  
 /tissue\_type="Embryonic Stem Cell"  
 /cell\_line="129.3 ES cells"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse Embryonic Stem (ES) cell (Lif+, 48 h, high density) cDNA library (Long)"  
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site.1: SalI; Site.2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]]. ES cells were plated at density 3x10<sup>4</sup>/cm<sup>2</sup>, on gelatin-coated plates and cultured for 48 hrs at 37 °C, 5% CO<sub>2</sub>. Culture medium: DMEM supplemented with 15% FBS, 2 mM L-glutamine, 0.1 mM NEAA, 1mM Sodium pyruvate, 0.1 mM beta-mercaptoethanol, 1000 U/ml LIF, 100 U/ml penicillin, and 100 U/ml streptomycin. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGAGCGCCGCTTTT-3'] from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker Lp-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.7 kb. The library was constructed by Yulan Piao."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 7; Length 252;  
 Best Local Similarity 94.4%; Pred. No. 5.9e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTCCAT 18  
 Db 127 CGACTGTGACTCTCCAT 110

RESULT 31  
 BG062707/c 266 bp mRNA linear EST 10-JUN-2003  
 LOCUS  
 DEFINITION  
 accession  
 version  
 keywords  
 organism  
 source  
 Mus musculus (house mouse)  
 Mus musculus  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 266)  
 Piao, Y., Ko, N.-T., Lim, M.-K. and Ko, M.-S. H.  
 Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

# JOURNAL

Genome Res. 11 (9), 1553-1558 (2001)  
 MEDLINE  
 PUBMED  
 11544199  
 On Jan 25, 2001 this sequence version replaced gi:12533453.  
 Other\_ESTs: U0956G04-3  
 Contact: George J. Kargul  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: [cdaa@lgsun.grc.nia.nih.gov](mailto:cdaa@lgsun.grc.nia.nih.gov)  
 nlaEST (<http://lgsun.grc.nia.nih.gov/cDNA/cDNA.html>)  
 Plate: U0956 row: G column: 04  
 Seq primer: -21M13 Reverse  
 High quality sequence stop: 266  
 POLYA-No.

## FEATURES

source  
 1. 266  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="nlaEST:U0956G04-5"  
 /db\_xref="taxon:10090"  
 /clone="L0956G04"  
 /tissue\_type="Newborn Kidney"  
 /dev\_stage="Newborn"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse Newborn Kidney cDNA library2 (Short)"  
 /note="Vector: pSPORT1 (Invitrogen); Site.1: SalI; Site.2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a short-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]]. In brief, double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGAGCGCCGCTTTT-3'] from 26 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker Lp-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-L. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 1.5 kb. The library was constructed by Yulan Piao(NIA)."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 4; Length 266;  
 Best Local Similarity 94.4%; Pred. No. 5.9e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTCCAT 18  
 Db 123 CGACTGTGACTCTCCAT 106

RESULT 32  
 CG668669/c 282 bp DNA linear GSS 02-OCT-2003  
 LOCUS  
 DEFINITION  
 accession  
 version  
 keywords  
 organism  
 source  
 Mus musculus (house mouse)  
 Mus musculus  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 282)  
 Piao, Y., Ko, N.-T., Lim, M.-K. and Ko, M.-S. H.  
 Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

REFERENCE  
 AUTHORS  
 TITLE

REFERENCE  
 AUTHORS  
 TITLE

REFERENCE 1 (bases 1 to 282)  
 AUTHORS Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, J., Finch, R.A.,  
 Figgott, J., Beltranda-Rio, H., Buxton, E.C., Edwards, J., Finch, R.A.,  
 Fiddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, M., Jaiting, C.,  
 Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,  
 Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,  
 Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,  
 Zhu, O., Person, C. and Sande, A.T.  
 TITLE Wnt1 kinase deficiency lowers blood pressure in mice: a gene-trap  
 screen to identify potential targets for therapeutic intervention  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
 COMMENT Contact: Zambrowicz BP  
 OmniBank  
 Lexicon Genetics Incorporated  
 4000 Research Forest Drive, The Woodlands, TX 77381, USA  
 Email: materials@lexgen.com  
 Gene trap sequence tag generated by 3' RACE from mouse ES cells as  
 described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)  
 Class: Gene Trap.  
 FEATURES  
 source Location/Qualifiers  
 1..282  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129Sv/Ev"  
 /db\_xref="taxon:10090"  
 /clone="OSF464816"  
 /cell\_type="embryonic stem cell"  
 /clone\_lib="Mus musculus 129Sv/Ev"  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 9; Length 282;  
 Best Local Similarity 94.4%; Pred. No. 6e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTGATCTCCAT 18  
 |||||  
 30 CGACTGTGATCTCCAT 13  
 |||||  
 RESULT 33 332 bp mRNA linear EST 11-DEC-2002  
 BY323607/c LOCUS  
 DEFINITION Musculus CDNA clone l030020D04 5', mRNA sequence.  
 VERSION BY323607  
 KEYWORDS  
 ACCESSION BY323607.1 GI:26514011  
 EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 332)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Mikaito, I., Osato, N., Saito, R., Suzuki, H., Yamana, I.,  
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
 Schönbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
 Home, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
 Choicha, C., Corradi, L.E., Cousins, S., Daila, E., Dragani, T.A.,  
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
 Kurochkin, I.V., Lee, Y., Lennard, B., Lyons, P.A., Maglott, D.R.,  
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,  
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
 Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,  
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
 Vetraro, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
 Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I.,  
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
 Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
 Rogers, U., Birney, E. and Hayashizaki, Y.  
 TITLE Analysis of the mouse transcriptome based on functional annotation  
 JOURNAL Nature 420, 563-573 (2002)  
 MEDLINE 22354683  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel.: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,  
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
 Shiraki, T., Tagami, M., Waki, K., Watabiki, A., Watanabe, M. and  
 Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome, 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Tissues were provided by Vasilius Aidinis (Biomedical Sciences  
 Research Center "Al. Fleming" Institute of Immunology 14-16 Al.  
 Fleming street 16672 Vardi, Greece) whose assistance we gratefully  
 acknowledge.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.  
 FEATURES  
 source Location/Qualifiers  
 1..332  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="L030020D04"  
 /cell\_type="synovial fibroblasts"  
 /clone\_lib="RIKEN full-length enriched, synovial  
 fibroblasts"  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 5; Length 332;  
 Best Local Similarity 94.4%; Pred. No. 6.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTGATCTCCAT 18  
 |||||  
 141 CGACTGTGATCTCCAT 124  
 |||||  
 RESULT 34 340 bp mRNA linear EST 06-DEC-2002  
 BY021968/c LOCUS  
 DEFINITION RIKEN full-length enriched, mammary gland RCB-0526  
 UY-MC(A) cDNA Mus musculus CDNA clone G830050F09 5', mRNA  
 sequence.

ACCESSION BY021968  
 VERSION BY021968.1 GI:26127411  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 340)  
 REFERENCE Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaio, I., Otsu, N., Saito, R., Suzuki, H., Yamana, K., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chottha, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reid, D. J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E., and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 JOURNAL MEDLINE 22354683  
 PUBLISHED 12466851  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numata, K., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Watanabe, M., and Hayashizaki, Y. Direct Substitution  
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.  
 Location/Qualifiers

source 1..340  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="G830050P09"  
 /tissue\_type="mammary gland"  
 /cell\_line="RCB-0526 JYG-MC(A)"  
 /clone\_lib="RIKEN full-length enriched, mammary gland RCB-0526 JYG-MC(A) cDNA"  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 5; Length 340;  
 Best Local Similarity 94.4%; Pred. No. 6.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 CGACTGTCATCTTCAT 18  
 Db 171 CGACTGTCATCTTCAT 154  
 RESULT 35  
 BY010085/c  
 LOCUS BY010085 351 bp mRNA linear EST 06-DEC-2002  
 DEFINITION BY010085 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus musculus cDNA clone G730016014 5', mRNA sequence.  
 ACCESSION BY010085  
 VERSION BY010085.1 GI:26070334  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 351)  
 REFERENCE Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaio, I., Otsu, N., Saito, R., Suzuki, H., Yamana, K., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chottha, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reid, D. J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E., and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 JOURNAL MEDLINE 22354683  
 PUBLISHED 12466851  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submision  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

## SOURCE

Location/Qualifiers  
1. 351  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="G730016014"  
/tissue\_type="lung"  
/cell\_line="RCB-0558 LLC"  
/clone\_lib="RIKEN full-length enriched, lung RCB-0558 LLC cDNA"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 351;  
Best Local Similarity 94.4%; Pred. No. 6.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CGACTGTGAATCTCCAT 18  
172 CGACTGTGAATCTCCAT 155

RESULT 36  
BY192197/c  
LOCUS  
DEFINITION  
BY192197 354 bp mRNA linear EST 10-DEC-2002  
BY192197 RIKEN full-length enriched, NOD-derived CD11c +ve dendritic cells Mus musculus cDNA clone F630326B04 5', mRNA sequence.  
VERSION  
BY192197  
KEYWORDS  
BY192197.1 GI:26368080  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 354)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishikido, T., Ohtsuka, N., Saito, R., Suzuki, H., Yamana, K., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schramm, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusio, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Geroldi, M., Gissi, C., Godzik, A., Gough, V., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, P., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,

## FEATURES

## SOURCE

Location/Qualifiers  
1. 354  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="taxon:10090"  
/clone="F630326B04"  
/cell\_line="RIKEN full-length enriched, NOD-derived CD11c +ve dendritic cells"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 354;  
Best Local Similarity 94.4%; Pred. No. 6.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CGACTGTGAATCTCCAT 18  
149 CGACTGTGAATCTCCAT 132

## COMMENT

Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Seiple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wysshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shigenaga, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
MEDLINE  
PUBMED  
22354683  
12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,

Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submision

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics

Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome

Trust/MRC building Addenbrookes Hospital Cambridge) whose

assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

RESULT 37  
LOCUS D76778/c  
DEFINITION MUS74C11 mouse embryonal carcinoma cell line F9 Mus musculus cDNA  
clone 74C11, mRNA sequence.  
ACCESSION D76778  
VERSION D76778.1 GI:1596448  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 354)  
AUTHORS Nishiguchi,S., Sakuma,R., Nomura,M., Zou,Z., Jearamakulwong,J., Jor,T., Yasunaga,T. and Shimada,K.  
TITLE A catalogue of genes in mouse embryonal carcinoma F9 cells identified with expressed sequence tags  
JOURNAL J. Biochem. 119 (4), 749-767 (1996)  
MEDLINE 96337530  
PUBMED 8743579  
COMMENT Contact: Kazunori Shimada  
Department of Medical Genetics, Division of Molecular Biomedicine  
Research Institute for Microbial Diseases, Osaka University  
3-1, Yamadaoka, Suita, Osaka, 565, Japan  
Tel: 06-879-8325  
Fax: 06-879-8326  
Location/Qualifiers  
1. .354  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="74C11"  
/clone\_lib="mouse embryonal carcinoma cell line F9"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 7; Length 354;  
Best Local Similarity 94.4%; Pred. No. 6.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 CGACTGTGATCTCTCCAT 18  
|||||  
106 CGACTGTGATCTCTCCAT 89  
|||||

RESULT 38  
LOCUS BY043120/c  
DEFINITION BY043120 397 bp mRNA linear EST 06-DEC-2002  
cell\_line=TRB-55BB88, etc. Mus musculus cDNA clone 1730027D23 5',  
mRNA sequence.  
ACCESSION BY043120  
VERSION BY043120.1 GI:26148563  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 397)  
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nishido,I., Otsu,N., Saito,R., Suzuki,H., Yamana,T., Kiyosawa,H., Yagi,K., Tomaru,Y., Haasegawa,T., Nogami,A., Schönbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schiraldi,M., Kanapin,A., Matsuda,H., Balalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V., Chothea,C., Corbani,L.R., Cousins,S., Dalla,B., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Guenichon,S., Hirokawa,N., Jackson,I.J., Jarvis,B.D., Kanai,A., Kawaji,H., Kawaawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglocz,D.R., Maltais,L., Marchionni,L., McKenzie,L., Mikti,H., Nagashima,T.,

Nunata,K., Okido,T., Pavan,W.J., Perlea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramchandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Secou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wyszewski-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavoian,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Kono,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.  
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
MEDLINE 22354683  
PUBMED 12466851  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watanabe,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RIKA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.  
Location/Qualifiers  
1. .397  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="1730027D23"  
/clone\_lib="RIKEN full-length enriched, pooled tissues, cell\_line=TRB-55BB88, etc."  
/note="Pooled tissues: (sex=mix, cell\_line=TRB-55BB88), (sex=mix, cell\_line=CRL-1722 L5178Y-R)"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 5; Length 397;  
Best Local Similarity 94.4%; Pred. No. 6.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 CGACTGTGATCTCTCCAT 18  
|||||  
Db 150 CGACTGTGATCTCTCCAT 133  
|||||

RESULT 39

CK577863/c  
 LOCUS CK577863 407 bp mRNA linear EST 16-JAN-2004  
 DEFINITION 1st WTS\_18124 AD-wrmcDNA library Caenorhabditis elegans cDNA 5' similar to F38A5.3, mRNA sequence.  
 ACCESSION CK577863  
 VERSION CK577863.1 GI:40961531  
 KEYWORDS EST.  
 SOURCE Caenorhabditis elegans  
 ORGANISM Caenorhabditis elegans  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodermidae; Caenorhabditis.  
 REFERENCE 1 (bases 1 to 407)  
 Li, S., Armstrong, C.M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P.O., Han, J.D., Chesneau, A., Hao, T., Goldberg, D.S., Li, N., Martineau, M., Raul, J.F., Lamesch, P., Xu, L., Tewari, M., Wong, S.L., Zhang, L.V., Bertiz, G.F., Jacotot, L., Vagstad, P., Redoul, O., Hitozane-Kishikawa, T., Li, Q., Gabel, H.W., Elewa, A., Baumgartner, B., Rose, D.J., Yu, H., Bosak, S., Segueria, R., Fraser, A., Mango, S.E., Saxton, W.M., Strom, S., Van Den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K.C., Harper, J.W., Cusick, M.E., Roth, F.P., Hill, D.E. and Vidal, M.  
 A Map of the Interactome Network of the Metazoan C. elegans Science (2004) In press  
 CONTACT: Vidal M  
 Marc Vidal Laboratory  
 Dana Farber Cancer Institute  
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
 Tel: 617 632 5180  
 Fax: 617 632 5739  
 Email: Marc.Vidal@dfci.harvard.edu  
 For the purpose of protein interaction mapping, we generated a C. elegans cDNA library (AD-wrmcDNA) in which poly(dT)-primed reverse transcribed cDNA are fused to the AD-encoding sequence of the yeast transcription factor GAL4. cDNAs were generated and cloned into the two hybrid vector pPC86 This interacting Sequence Tag 1st WTS\_18124 (F38A5.3) interacts as a prey with the bait F08G2.5  
 PCR primers  
 FORWARD: CGCGTTGGAATCACTACAGGG  
 BACKWARD: CGAGACTGACCAACTCTGGCG  
 Insert Length: 407 Std Error: 47.00  
 Plate: 227 row: 02 column: F  
 Seq primer: CGCGTTGGAATCACTACAGGG  
 High quality sequence stop: 406  
 POLYA-No.  
 FEATURES  
 SOURCE location/Qualifiers  
 1..407  
 /organism="Caenorhabditis elegans"  
 /mol\_type="mRNA"  
 /strain="N2"  
 /db\_xref="taxon:6239"  
 /set="male, hermaphrodite"  
 /dev\_stage="embryo, L1, L2, L3, L4, adult, dauer"  
 /clone\_lib="AD-wrmcDNA library"  
 /note="Vector: pPC86; For the purpose of protein interaction mapping, we generated a C. elegans cDNA library (AD-wrmcDNA) in which poly(dT)-primed reverse transcribed cDNA are fused to the AD-encoding sequence of the yeast transcription factor GAL4. This library was made with poly(A) + RNA isolated from mated populations of wild-type (N2 strain) animals of all stages of development including embryonic, larval (L1 to L4 stages), adults and dauer. Approximately equal quantities of RNA from different populations were acquired. cDNAs were generated and cloned into the two hybrid vector pPC86. The library contains ~3\*10e7 clones. Reference - GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFsomes - Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lironi MA, van den Heuvel S, Vidal M - Methods Enzymol. 2000;328:575-92"

Best Local Similarity 94.4%; Pred. No. 6.3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 CGACTGGAATCTCCAT 18  
 Db 101 CGACTGGAATCTCCCT 84  
 RESULT 40  
 AM209975/c 428 bp mRNA linear EST 03-DEC-1999  
 LOCUS u149d12.y1 Rasbass mouse MOV 9 5 optic vesicle Mus musculus cDNA  
 DEFINITION clone IMAGE:2101655 5' similar to gb:X01237 Mouse mRNA for cellular tumour antigen p53 (MOUSE);, mRNA sequence.  
 ACCESSION AM209975 GI:6515915  
 VERSION AM209975.1 GI:6515915  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 428)  
 Marra, M., Hallier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 The WashU-NCI Mouse EST Project 1999  
 Unpublished (1999)  
 CONTACT: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:998587  
 Seq primer: primer name ambiguous.  
 FEATURES  
 SOURCE location/Qualifiers  
 1..428  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CD-1"  
 /db\_xref="taxon:10090"  
 /tissue\_type="optic vesicle and lens placode"  
 /dev\_stage="embryo - 9.5 dpc"  
 /lab\_host="DH10B"  
 /clone\_lib="Rasbass mouse MOV 9 5 optic vesicle"  
 /note="Organ: eye; Vector: pSPORT1; mRNA made from developing eye tissue, cDNA made by oligo-dT priming with NotI oligo. Salt adaptor (5'-TCGACCGACGCGCG-3') ligated to 5' ends. Size-selected with cDNA size fractionation resin, average insert size 1.3 kb. Primary library, non-amplified. Library constructed by Dr. Pen Rasbass (penr@gu.mrc.ac.uk)."  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 2; Length 428;  
 Best Local Similarity 94.4%; Pred. No. 6.3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 CGACTGGAATCTCCAT 18  
 Db 125 CGACTGGAATCTCCAT 108  
 RESULT 41  
 BB846564/c 430 bp mRNA linear EST 26-NOV-2001  
 LOCUS BB846564 RIKEN full-length enriched, adult male kidney Mus musculus  
 DEFINITION cDNA clone F530003D19 5', mRNA sequence.

**ACCESSION** BB846564  
**VERSION** BB846564.1 GI:17084939  
**KEYWORDS** EST.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
**REFERENCE** 1 (bases 1 to 430)  
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imocani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Komoto,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahata,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A., Miyamatsu,M. and Hayashizaki,Y.  
**TITLE** RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)  
**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komoto,H., Okazaki,Y., Miyamatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 waga,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Miyamatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Komoto,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 further detail.  
 e mouse tissues.  
**FEATURES** Location/Qualifiers  
 source  
 1. 430  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="F530003D19"  
 /sex="male"  
 /tissue\_type="kidney"  
 /dev\_stage="adult"  
 /lab\_host="SOLR"  
 /clone\_lib="RIKEN full-length enriched, adult male kidney"  
 /note="Site 1: XhoI; Site 2: SclI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'  
 GAGGAGAGAGCGCGCCCACTGATTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'  
 GAGGAGAGAGATCCAGATCAATTAATTAATTAATTAACCCCCCCCC 3'].  
 cDNA was cleaved with XhoI and SclI."

**ORIGIN**  
 Query Match 91.1%; Score 16.4; DB 2; Length 430;  
 Best Local Similarity 94.4%; Pred. No. 6.3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 1 CGACTGTAATCCCTCAT 18  
 141 CGACTGTAATCCCTCAT 124  
**RESULT** 42  
 CA895446/c  
**LOCUS** CA895446  
**DEFINITION** CA895446 447 bp mRNA linear EST 20-DEC-2002  
 Library (long) Mus musculus cDNA clone NIA:B0192803 IMAGE:30102638  
 5', mRNA sequence.  
**ACCESSION** CA895446  
**VERSION** CA895446  
**KEYWORDS** EST.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
**REFERENCE** 1 (bases 1 to 447)  
 Piao,Y., Dudekula,D.B., Qian,Y., Martin,P.R., Alpa,K., Vescovi,A.L. and Ko,M.S.H.  
**TITLE** Systematic Analyses of NIA Mouse Neural Stem Cell (Differentiated) cDNA Library (Long)  
**JOURNAL** Unpublished (2002)  
**COMMENT** Contact: Dawood B. Dudekula  
 Laboratory of Gene  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@gsun.gsc.nia.nih.gov  
 Plate: B0192 row: B column: 03  
 Seq primer: -21M13 Reverse  
 High quality sequence stop: 447  
 POLYA=No.  
**FEATURES** Location/Qualifiers  
 source  
 1. 447  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CD1"  
 /db\_xref="nibEST:B0192803-5N"  
 /db\_xref="taxon:10090"  
 /clone="NIA:B0192803 IMAGE:30102638"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse Neural Stem Cell (Differentiated) cDNA Library (Long)"  
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetic, National Institute on Aging (NIA), Intramural Research Program, NIH (http://ligsun.gsc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). PMID: 11544191). Total RNAs were obtained from Dr. Angelo L. Vescovi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTGTAATCCAGATCCAGGCGCGCCCTTTTCTTTT-3'] from 2.0 microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Gentriicon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Gentriicon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about

3.2 kb. The library was constructed by Yulan Piao."

Query Match 91.1%; Score 16.4; DB 6; Length 447;  
Best Local Similarity 94.4%; Pred. No. 6.4e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
| | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | |  
Db 114 CGACTGTGAATCCTCCAT 97

RESULT 43  
CG501482/c 447 bp DNA linear GSS 01-OCT-2003  
LOCUS OST45353 Mus musculus 1295v/Er Mus musculus genomic clone OST45353.  
DEFINITION genomic survey sequence.

ACCESSION CG501482.1 GI:37275240

VERSION CG501482.1 GI:37275240

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 447)

AUTHORS Zambrowicz, B.P., Abulin, A., Ramirez-Solis, R., Richter, L.J.,  
Pisgott, J., Beltrandelio, H., Buxton, E.C., Edwards, J., Finch, R.A.,  
Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Jais, C.,  
Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,  
Payne, R., Porter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,  
Sparks, M.O., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,  
Zhu, Q., Person, C., and Sands, A.T.

TITLE Mml kinase deficiency lowers blood pressure in mice: a gene-trap  
screen to identify potential targets for therapeutic intervention

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

COMMENT Contact: Zambrowicz BP

OmniBank

Lexicon Genetics Incorporated

4000 Research Forest Drive, The Woodlands, TX 77381, USA

Email: materials@lexgen.com

Gene trap sequence tag generated by 3' RACE from mouse ES cells as

described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)

Class: Gene Trap.

Location/Qualifiers

1. .447

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="1295v/Er"

/db\_xref="taxon:10090"

/clone="OST45353"

/cell\_type="embryonic stem cell"

/clone\_id="Mus musculus 1295v/Er"

ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 447;  
Best Local Similarity 94.4%; Pred. No. 6.4e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
| | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | |  
Db 34 CGACTGTGAATCCTCCAT 17

RESULT 44  
BB858888 RIKEN full-length enriched, Nullipotent stem cell CRJ-2070  
LOCUS BB858888 RIKEN full-length enriched, Nullipotent stem cell CRJ-2070  
DEFINITION NE cDNA Mus musculus CDNA clone G43003B20 5', mRNA sequence.

ACCESSION BB858888.1 GI:17100342

VERSION BB858888.1 GI:17100342

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 462)

Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Akimura, T., Arawaka, T., Carninci, P., Furuno, M., Hanagaki, T.,  
Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K.,  
Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,  
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Tanaka, T., Tomaru, A., Toyota, T., Watanabe, A., Yasunishi, A.,  
Muramatsu, M., and Hayashizaki, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)

wag, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.,  
and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multichannel sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

e mouse tissues.

Location/Qualifiers

1. .462

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C3H/J"

/db\_xref="taxon:10090"

/clone="G43003B20"

/cell\_type="Nullipotent stem cell"

/cell\_line="CRJ-2070 NE"

/dev\_stage="7 days embryo"

/clone\_id="RIKEN full-length enriched, Nullipotent stem  
cell CRJ-2070 NE cDNA"

ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 462;  
Best Local Similarity 94.4%; Pred. No. 6.4e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
| | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | |  
Db 144 CGACTGTGAATCCTCCAT 127

RESULT 45  
CF726263/466 bp mRNA linear EST 09-OCT-2003  
LOCUS CF726263/466 bp mRNA linear EST 09-OCT-2003  
DEFINITION UI-M-GZO-cjm-d-20-0-UI r1 NIH BMP\_GZO Mus musculus cDNA clone  
IMAGE:30603259 5', mRNA sequence.

ACCESSION CF726263

VERSION CF726263

CF726263.1 GI:37600431



**KEYWORDS**  
**SOURCE**  
**ORGANISM**  
**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**  
**COMMENT**

EST.  
 Mus musculus (house mouse)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 466)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. James Lin University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mouse1.html>  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
 Seg primer: PYX-5.  
 Location/Qualifiers  
 1..466  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:30603259"  
 /tissue\_type="whole eye"  
 /dev\_stage="embryo 12.5,13.5,14.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH BMAP G20"  
 /note="Organ: Eye; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTTATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., Program coordinator."

**ORIGIN**

Query Match 91.1%; Score 16.4; DB 7; Length 466;  
 Best Local Similarity 94.4%; Pred. No. 6.4e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCCAT 18  
 |||||  
 Db 231 CGACTGTGATCTCCAT 214

**RESULT 46**  
**CAS33945/c**  
**LOCUS**  
**DEFINITION**  
**ACCESSION**  
**VERSION**  
**KEYWORDS**  
**SOURCE**  
**ORGANISM**  
**REFERENCE**  
**AUTHORS**  
**TITLE**

CAS33945 486 bp mRNA linear EST 18-NOV-2002  
 C0409809-5N NIA Mouse ES Cell (LIF-) cDNA library (Long) Mus  
 musculus cDNA clone NIA:C0409809 IMAGE:30010484 5', mRNA sequence.  
 CAS33945  
 EST.  
 GI:25065602  
 Mus musculus (house mouse)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 486)  
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K.,  
 Jaradit, S.A., Boheler, K.R. and Ko, M.S.H.  
 Systematic Analyses of NIA Mouse ES Cell (LIF-) cDNA library (Long)

**JOURNAL**  
**COMMENT**

Unpublished (2001)  
 Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: [cdna@lgsun.grc.nia.nih.gov](mailto:cdna@lgsun.grc.nia.nih.gov)  
 Plate: C0409 row: B column: 09  
 Seg primer: M13 Reverse  
 High quality sequence stop: 486  
 POLYA=No.

**FEATURES**  
**SOURCE**

Location/Qualifiers  
 1..486  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129/Sv x 129/Sv-CP"  
 /db\_xref="EAST:C0409809-5N"  
 /db\_xref="taxon:10090"  
 /clone="NIA:C0409809 IMAGE:30010484"  
 /tissue\_type="ES Cell (LIF-)"  
 /cell\_line="F1 ES cells"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse ES Cell (LIF-) cDNA library (Long)"  
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). PMID: 11544193). Total RNAs were obtained from Dr. Kenneth R. Boheler (National Institute on Aging, USA). ES cells were cultured without feeder cells in the absence of LIF for 4 hours or 18 hours. Equimolar mixture of these RNA samples was used for the library construction. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen: 5'-TGACTGTCTGATCGGAGCGCCCTTTTCTTTT-3') from 21 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker l1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao (NIA)."

**ORIGIN**

Query Match 91.1%; Score 16.4; DB 6; Length 486;  
 Best Local Similarity 94.4%; Pred. No. 6.4e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCCAT 18  
 |||||  
 Db 121 CGACTGTGATCTCCAT 104

**RESULT 47**  
**CAS65317/c**  
**LOCUS**  
**DEFINITION**  
**ACCESSION**  
**VERSION**  
**KEYWORDS**  
**SOURCE**  
**ORGANISM**  
**REFERENCE**  
**AUTHORS**  
**TITLE**

CAS65317 506 bp mRNA linear EST 19-NOV-2002  
 K0336E11-5N NIA Mouse Osteoblast cDNA library (Long) Mus musculus  
 cDNA clone NIA:K0336E11 IMAGE:30057754 5', mRNA sequence.  
 CAS65317  
 EST.  
 GI:25109988  
 Mus musculus (house mouse)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 506)  
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,

TITLE  
Journal  
COMMENT  
Umezawa, A. and Ko, M.S.H.  
Systematic Analyses of NIA Mouse Osteoblast cDNA Library (Long)  
Unpublished (2001)  
Other ESTs: K0336E11-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov  
Plate: K0336 row: E column: 11  
Seq primer: M13 Reverse  
High quality sequence stop: 506  
POLYA=No.

#### FEATURES

source

Location/Qualifiers  
1..506  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C3H/He"  
/db\_xref="taeST:K0336E11-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0336E11 IMAGE:30057754"  
/tissue\_type="Osteoblast"  
/cell\_line="XUSA-A1 cells"  
/lab\_host="DH10B"  
/clone\_id="NIA Mouse Osteoblast cDNA Library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were obtained from Dr. Akhiro Umezawa (Keio University School of Medicine, Japan). Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-GACAGCTGATCGAGCGCGCCCTTTT-3'] from 2.1 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao (NIA)."

#### ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 506;  
Best Local Similarity 94.4%; Pred. No. 6.5e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTCGATCCTCCAT 18  
Db 111 CGACTGTCGATCCTCCAT 94

RESULT 48  
CG593379/c 528 bp DNA linear GSS 02-OCT-2003  
LOCUS OST250499 Mus musculus 129SV/Er Mus musculus genomic clone  
DEFINITION OST250499, genomic survey sequence.  
ACCESSION CG593379  
VERSION CG593379.1 GI:37403364  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 528)  
Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J.,

AUTHORS

TITLE  
Journal  
COMMENT  
Piggott, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A.,  
Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jang, C.,  
Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-O., Markesich, D.,  
Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,  
Spark, M.J., Van Sigtendijk, I., Vogel, P., Walke, W., Xu, N.,  
Zhu, Q., Pearson, C. and Sands, A.T.  
Mx1 kinase deficiency lowers blood pressure in mice: a gene-trap  
screen to identify potential targets for therapeutic intervention  
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
Contact: Zambrowicz BP  
OmniBank  
Lexicon Genetics Incorporated  
4000 Research Forest Drive, The Woodlands, TX 77381, USA  
Email: materials@lexgen.com  
Gene trap sequence tag generated by 3' RACE from mouse ES cells as  
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)  
Class: Gene trap.

#### FEATURES

source

Location/Qualifiers  
1..528  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129SV/Er"  
/db\_xref="taxon:10090"  
/db\_xref="OST250499"  
/cell\_type="embryonic stem cell"  
/clone\_id="Mus musculus 129SV/Er"

#### ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 528;  
Best Local Similarity 94.4%; Pred. No. 6.5e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTCGATCCTCCAT 18  
Db 35 CGACTGTCGATCCTCCAT 18

RESULT 49  
CD560905 529 bp mRNA linear EST 11-JUN-2003  
LOCUS B0421A08-5 NIA Mouse E6.5 Whole Embryo cDNA Library (Long) Mus  
musculus cDNA clone NIA:B0421A08 IMAGE:3040647 5', mRNA sequence.  
ACCESSION CD560905  
VERSION CD560905.1 GI:31604798  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 529)  
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
Construction of long-transcript enriched cDNA libraries from  
submicrogram amounts of total RNAs by a universal PCR amplification  
method  
Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL MEDLINE 21429098  
PUBMED 11544199  
COMMENT Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov  
Plate: B0421 row: A column: 08  
Seq primer: M13 Reverse  
High quality sequence stop: 529  
POLYA=No.

#### FEATURES

source

Location/Qualifiers  
1..529  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taeST:B0421A08-5"  
/db\_xref="taxon:10090"

/clone="N1A:80421A08 IMAGE:30440647"  
 /tissue\_type="E6.5 whole embryo"/  
 /dev\_stage="whole embryo including extraembryonic tissues  
 at 6.5-days postcoitum"  
 /lab\_host="DH10B"  
 /clone\_11b="N1A Mouse E6.5 Whole Embryo cDNA Library  
 (long)"  
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site\_1: SalI;  
 Site\_2: NotI; Mouse cDNA project by the Laboratory of  
 Genetics, National Institute on Aging (NIA), Intramural  
 Research Program, NIH (<http://lglgen.grc.nia.nih.gov/cDNA>).  
 This is a long-transcript enriched cDNA library (Ref.  
 Genome Res. 11: 1553-1558 (2001)). [PMID: 11544191]. Total  
 RNAs were extracted from a pool of 7 embryos at 6.5-days  
 postcoitum. Double-stranded cDNAs were synthesized with an  
 Oligo(dT) primer [Invitrogen:  
 5'-GACATGATTCAAGACCGACGCGCCCTTTTTTTTTTTTTT-3']  
 from 0.53 ug of total RNA, treated with T4 DNA polymerase,  
 and purified by ethanol-precipitation. The cDNAs were  
 ligated to Lone-linker Lr-Sal4, purified by  
 phenol/chloroform, and separated from free linkers by  
 centrifugation 100. Then, the cDNAs were amplified by  
 long-range high fidelity PCR using Ex Taq polymerase  
 (Takara) with a primer Sal4-S. The products were purified  
 by phenol/chloroform and centrifugation 100. The cDNAs were  
 digested with SalI and NotI enzymes and cloned into  
 SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E.  
 coli host was transformed with the ligation mixture by the  
 standard chemical method. The average insert size is about  
 2.3kb. The library was constructed by TulaN Piao."

**ORIGIN**

Query Match	91.1%	Score 16.4	DB 6	Length 529
Best Local Similarity	94.4%	Pred. No. 6.5e+02		
Matches 17	Conservative	0	Mismatches 1	Indels 0
Qy	1	CGACGTGATCTCTCCAT	18	
Db	120	CGACGTGATCTCTCCAT	103	

RESULT	50
CAS33770/c	
LOCUS	530 bp mRNA linear EST 18-NOV-2007
DEFINITION	CA046D02-SN NIA Mouse ES Cell (LIF-) cDNA library (Long) Mus musculus cDNA clone NIA: C0406D02 IMAGE:30010213 5' mRNA sequence.
ACCESSION	CA533770
VERSION	CA533770.1 GI:25065205
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Mamaliyotai, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 530) Pilo, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaraedt, S.A., Boheler K.R. and Ko, M.S.H. Systematic Analyses of NIA Mouse ES Cell (LIF-) cDNA Library (long) Unpublished (2001)
TITLE	Contact: Dawood B. Dudekula Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdha@gsun-grc.nia.nih.gov
JOURNAL	Plate: C0406 row: D column: 02 Seq primer: M13 Reverse High quality sequence stop: 530 POLYA-No.
COMMENT	

```
FEATURES
SOURCE
location/Qualifiers
1..530
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x 129/Sv-CP"
/db_xref="niasST:C0406D02-5N"
```

/db\_xref=taxon:10090"  
 /clone=NIA:CG406CD2 IMAGE:30010213"  
 /cruise\_type="ES Cell (LIF-)"  
 /cell\_line="R1 ES cells"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH Mouse ES Cell (LIF-) cDNA Library (Long)"  
 /note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
 NotI; Mouse cDNA project by the Laboratory of Genetics,  
 National Institute on Aging (NIA), Intramural Research  
 Program, NIH (<http://igsn.gsc.nia.nih.gov/cDNA>). This is  
 a long-transcript enriched cDNA library (Ref. Genome Res.  
 11: 1553-1558 (2001). [PMID 11544199]). Total RNAs were  
 obtained from Dr. Kenneth R. Boheler (National Institute  
 on Aging, USA). ES cells were cultured without feeder  
 cells in the absence of LIF for 4 hours or 18 hours.  
 Aqueous mixture of these RNA samples was used for the  
 library construction. Double-stranded cDNAs were  
 synthesized with an oligo(dT) primer (Invitrogen):  
 5'-GCACGAGTCGACGACGACGAGCGCCCTTTTCTTTTCTTTT-3' from  
 21 ug of total RNA, treated with T4 DNA polymerase, and  
 purified by ethanol-precipitation. The cDNAs were ligated  
 to lone-linker *NotI*-SalI, purified by phenol/chloroform, and  
 separated from free linkers by Centricon 100. Then, the  
 cDNAs were amplified by long-range high fidelity PCR using  
 Ex Taq polymerase (Takara) with a primer SalI-S. The  
 products were purified by phenol/chloroform and Centricon  
 100. The cDNAs were digested with SalI and NotI enzymes  
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
 The DH10B E. coli host was transformed with the ligation  
 mixture by the standard chemical method. The average  
 insert size is about 2.4 kb. The library was constructed  
 by Yulan Flao (NIA).".

**ORIGIN**

Query Match	91.1%	Score 16.4	DB 6	Length 530
Best Local Similarity	94.4%	Pred. No. 6.5e+02		
Matches 17; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0
QY	1	CGACTGTCGATCCTCCAT	18	
Db	121	CGACTGTCGATCCTCCAT	104	

Search completed: January 14, 2005, 17:37:54  
Job time : 2450.87 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:09:04 ; Search time 301.263 Seconds  
(without alignments)  
313.645 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18  
Sequence: 1 CGACTGGAATCCCTCCAT 18

Scoring table: IDENTITY\_NUC  
Gapc 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 500 summaries

Database :

N\_Geneseq\_23Sep04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002as:\*  
7: Geneseq2002bs:\*  
8: Geneseq2003as:\*  
9: Geneseq2003bs:\*  
10: Geneseq2003cs:\*  
11: Geneseq2003ds:\*  
12: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	2	AAQ88280 Antisense
2	18	100.0	1416	3	AAZ52303 Rat p35 (
3	18	100.0	1627	6	ABK63584 Rat seque
4	18	100.0	1627	12	ADP72925 Renal tox
5	18	100.0	1786	10	ADBS5550 Toxicity-
6	18	100.0	1786	10	ADBS5157 Primary r
7	16.4	91.1	93	12	ADL71108 p53 hairp
8	16.4	91.1	390	6	AAQ67682 Bacillus
9	16.4	91.1	1173	2	AAQ67683 Murine p5
10	16.4	91.1	1173	2	AAQ67683 Murine p5
11	16.4	91.1	1208	12	ADJ32551 Mouse p53
12	16.4	91.1	1208	12	ADJ32551 Mouse p53
13	15.4	85.6	525	2	AAQ85528 DNA probe
14	15.4	85.6	525	2	AAQ85528 DNA probe
15	15.4	85.6	584	5	AAQ85528 Human CDN
16	15.4	85.6	608	5	AAQ85528 Human CDN
17	15.4	85.6	795	12	ADN62510 A. thalia
18	15.4	85.6	806	3	AAQ85528 Human CDN
19	15.4	85.6	806	3	AAQ85528 Human CDN
20	15.4	85.6	870	4	AAQ85528 Human CDN
21	15.4	85.6	870	4	AAQ85528 Human CDN

22	15.4	85.6	870	12	ADL30326	Adl30326 3' end of
23	15.4	85.6	1077	3	AAQ40009	AAQ40009 Arabidops
24	15.4	85.6	1364	4	AAH14165	AAH14165 Human CDN
25	15.4	85.6	1487	10	ADDB9659	ADDB9659 Human REM
26	15.4	85.6	1490	4	AAQ94874	AAQ94874 Human ful
27	15.4	85.6	1490	12	ADL32031	Adl32031 Full leng
28	15.4	85.6	1510	5	AAH64732	AAH64732 Human sec
29	15.4	85.6	1523	5	AAH64771	AAH64771 Human sec
30	15.4	85.6	1523	5	AAV43615	AAV43615 Human sec
31	15.4	85.6	1664	8	ABZ68515	ABZ68515 Nucleotid
32	15.4	85.6	1664	12	ADP18659	ADP18659 Human TAT
33	15.4	85.6	1755	6	AAV55741	AAV55741 Human sec
34	15.4	85.6	1755	6	ABQ92067	ABQ92067 Human pol
35	15.4	85.6	2148	3	AAQ21834	AAQ21834 Human bre
36	15.4	85.6	2243	4	ABL19347	ABL19347 Drosophi
37	15.4	85.6	2243	4	ABL19347	ABL19347 Drosophi
38	15.4	85.6	3703	4	ABL02040	ABL02040 Drosophi
39	15.4	85.6	4357	4	AAK75168	AAK75168 Human imm
40	15.4	85.6	4532	4	ABL19346	ABL19346 Drosophi
41	15.4	85.6	6446	4	ABL15658	ABL15658 Drosophi
42	15.4	85.6	20000	12	ADQ47193	Ado47193 DNA seque
43	15	83.3	18	10	ACA61261	ACA61261 Rat p35 a
44	15	83.3	1024	3	AAZ97144	AAZ97144 Human pro
45	15	83.3	1065	6	ABN66254	ABN66254 Streptoco
46	15	83.3	11000	6	ABN71527	ABN71527 01
47	14.8	82.2	93	12	ADL71108	Adl71108 p53 hairp
48	14.8	82.2	281	12	ADQ05153	Adq05153 Soybean h
49	14.8	82.2	417	6	ABL66464	ABL66464 Lung can
50	14.8	82.2	417	6	ABL65765	ABL65765 Lung can
51	14.8	82.2	417	6	ABN93583	ABN93583 Gene #81
52	14.8	82.2	474	9	ACH32462	ACH32462 Human end
53	14.8	82.2	781	6	AAI45845	AAI45845 Human sec
54	14.8	82.2	791	4	AAI24530	AAI24530 Human bre
55	14.8	82.2	908	12	ADU42388	Adj42388 Plant CDN
56	14.8	82.2	966	4	ABA89005	ABA89005 Escherich
57	14.8	82.2	1002	4	AAQ59713	AAQ59713 Propionib
58	14.8	82.2	1002	8	ACF64642	ACF64642 Propionib
59	14.8	82.2	1167	5	AAH46482	AAH46482 Human rib
60	14.8	82.2	1170	4	AAH18499	AAH18499 Human CDN
61	14.8	82.2	1185	4	AAH14287	AAH14287 Human CDN
62	14.8	82.2	1185	12	ADQ020164	Ado20164 Human RNA
63	14.8	82.2	1188	6	AAQ24384	AAQ24384 Human RNA
64	14.8	82.2	2388	4	ABL23309	ABL23309 Drosophi
65	14.8	82.2	2489	6	ABQ78877	ABQ78877 E. coli C
66	14.8	82.2	2489	10	ADH80444	Adh80444 Escherich
67	14.8	82.2	2498	4	ABA89004	ABA89004 Escherich
68	14.8	82.2	2541	6	AAI45816	AAI45816 Human sec
69	14.8	82.2	2557	4	ABL29399	ABL29399 Drosophi
70	14.8	82.2	2832	4	ABL05208	ABL05208 Drosophi
71	14.8	82.2	4388	4	ABL23308	ABL23308 Drosophi
72	14.8	82.2	4681	4	ABL23398	ABL23398 Drosophi
73	14.8	82.2	4725	4	ABL23322	ABL23322 Drosophi
74	14.8	82.2	42035	8	AAQ59562	AAQ59562 Propionib
75	14.8	82.2	42035	8	AAQ59562	AAQ59562 Propionib
76	14.8	82.2	42035	8	AAQ59562	AAQ59562 Propionib
77	14.8	82.2	110000	6	ABX08336	ABX08336 11
78	14.8	82.2	110000	12	ADU25985	ADU25985 12
79	14.8	82.2	110000	12	ADU25985	ADU25985 12
80	14.8	82.2	110000	12	ADN97989	ADN97989 12
81	14.8	82.2	110000	12	ADN97989	ADN97989 12
82	14.8	82.2	110000	12	ADQ50281	ADQ50281 12
83	14.8	82.2	110000	12	ADQ50281	ADQ50281 12
84	14.4	80.0	35	12	ADN68991	Adn68991 Mouse p53
85	14.4	80.0	331	4	AAI20107	AAI20107 Probe #10
86	14.4	80.0	331	4	AAQ65136	AAQ65136 Human foe
87	14.4	80.0	331	4	AAI45307	AAI45307 Probe #13
88	14.4	80.0	331	4	AAQ47251	AAQ47251 Human bre
89	14.4	80.0	331	4	AAQ32338	AAQ32338 Probe #10
90	14.4	80.0	331	4	AAQ32338	AAQ32338 Human bon
91	14.4	80.0	331	4	AAK13555	AAK13555 Human bra
92	14.4	80.0	331	4	ABQ38883	ABQ38883 Human liv
93	14.4	80.0	331	5	AAI05813	AAI05813 Probe #58
94	14.4	80.0	331	6	ABQ13381	ABQ13381 Human gen

95	14.4	80.0	376	6	ABO85387	Arabidops	c 168	14.4	80.0	6911	2	AAT34177	Act34177	Corynefor
96	14.4	80.0	413	9	ACH20958	Human adu	169	14.4	80.0	8034	2	ABL03008	AbI03008	Drosophi1
97	14.4	80.0	427	8	AX38233	Bovine ES	c 170	14.4	80.0	8674	8	ABX63814	AbX63814	Human cdn
98	14.4	80.0	433	9	ACH21300	Human adu	c 171	14.4	80.0	8878	2	AAT03852	Act03852	Human fib
99	14.4	80.0	444	4	AA110868	Probe #80	c 172	14.4	80.0	8878	6	AAD45341	Aad45341	Human ant
100	14.4	80.0	444	4	ABA52519	Human foe	c 173	14.4	80.0	8878	9	AAD55117	Aad55117	Human ant
101	14.4	80.0	444	4	AA132128	Probe #81	c 174	14.4	80.0	8878	11	ADPE5553	Adpe5553	Human FBG
102	14.4	80.0	444	4	ABA42097	Human bre	c 175	14.4	80.0	8878	11	ADPE5553	Adpe5553	Human fib
103	14.4	80.0	444	4	ABA22308	Probe #77	c 176	14.4	80.0	11840	11	ADPE5553	Adpe5553	Human fib
104	14.4	80.0	444	4	AAK26235	Human bon	c 177	14.4	80.0	11840	11	AAK83820	Aak83820	Human imm
105	14.4	80.0	444	4	AAK00783	Human bra	c 178	14.4	80.0	32195	9	ABK42625	Abk42625	Genomic s
106	14.4	80.0	444	4	ABE25828	Human liv	c 179	14.4	80.0	39796	3	AAAC6161	Aac6161	Connectiv
107	14.4	80.0	444	5	AA100792	Human liv	c 180	14.4	80.0	110000	2	AAZ01425	Aaz01425	Nucleoid
108	14.4	80.0	445	6	ABSO0823	Probe #78	c 181	14.4	80.0	349980	5	AAH68531	Aah68531	Continuati
109	14.4	80.0	445	6	ABSO0823	Human gen	c 182	14.4	80.0	385	5	AAH68531	Aah68531	(7 of
110	14.4	80.0	455	3	AACT71409	Single nu	c 183	14.4	77.8	385	10	AAK55788	Aak55788	C glutami
111	14.4	80.0	455	3	AACT71418	Single nu	c 184	14.4	77.8	387	10	ACD97529	Acd97529	Human imm
112	14.4	80.0	477	9	ACH21912	Single nu	c 185	14.4	77.8	465	8	ABZ55728	Abz55728	Human col
113	14.4	80.0	495	6	ABG69615	Novel mur	c 186	14.4	77.8	529	10	ADK57503	Adk57503	Aspergill1
114	14.4	80.0	573	4	AAI17550	Probe #74	c 187	14.4	77.8	750	8	AAI24883	Aai24883	Plant DNA
115	14.4	80.0	573	4	ABA62483	Human foe	c 188	14.4	77.8	852	3	AAAC43068	Aac43068	Human bre
116	14.4	80.0	573	4	AAI142463	Probe #11	c 189	14.4	77.8	1136	3	AAAC8191	Aac8191	Arabidops
117	14.4	80.0	573	4	ABA29814	Probe #82	c 190	14.4	77.8	1138	3	AAAC5620	Aac5620	Arabidops
118	14.4	80.0	573	4	AAK36698	Human bon	c 191	14.4	77.8	1234	6	ABG65642	Abg65642	Mouse Mga
119	14.4	80.0	573	4	AAK10835	Human bra	c 192	14.4	77.8	1395	2	AAAI5662	Aai5662	Protein p
120	14.4	80.0	573	4	ABG36357	Human liv	c 193	14.4	77.8	1395	10	ADK64853	Adk64853	Disease t
121	14.4	80.0	573	6	ABSI0699	Human gen	c 194	14.4	77.8	2088	12	ADN74420	Adn74420	Thale cre
122	14.4	80.0	598	12	ACH73732	Human gen	c 195	14.4	77.8	2125	12	ADG20490	Adg20490	Mouse pro
123	14.4	80.0	864	6	ABZ14669	Arabidops	c 196	14.4	77.8	2430	3	AACT3972	Aac3972	Arabidops
124	14.4	80.0	864	12	ADDO1806	Thalecres	c 197	14.4	77.8	3109	12	ADJ87641	Adj87641	Nervous s
125	14.4	80.0	920	2	ABN98651	Arabidops	c 198	14.4	77.8	4914	8	ABV76029	Abv76029	Mouse pla
126	14.4	80.0	1093	2	AAT95761	Arabidops	c 199	14.4	77.8	13813	4	AAK89790	Aak89790	Human dig
127	14.4	80.0	1093	2	AAT95761	Arabidops	c 200	14.4	77.8	13904	10	ADG87172	Adg87172	Human GPC
128	14.4	80.0	1098	5	AAH67333	C glucami	c 201	14.4	77.8	42048	2	AAK71918	Aak71918	Human imm
129	14.4	80.0	1118	3	AAH67333	C glucami	c 202	14.4	77.8	43226	2	AAK60263	Aak60263	Human imm
130	14.4	80.0	1464	3	AAAC52804	Arabidops	c 203	13.8	76.7	25	9	ADH23526	Adh23526	Nucleic a
131	14.4	80.0	1629	12	ADB76867	Human cdn	c 204	13.8	76.7	34	9	ADCK15038	Adck15038	PCR prime
132	14.4	80.0	1850	2	AAV09251	Cyclochrom	c 205	13.8	76.7	34	10	ADA02250	Ada02250	Mouse car
133	14.4	80.0	1850	2	AAV12203	Zebrafish	c 206	13.8	76.7	49	10	ADH31989	Adh31989	Human mic
134	14.4	80.0	1850	2	AAH22442	Zebrafish	c 207	13.8	76.7	50	12	ADM32608	Adm32608	Mouse car
135	14.4	80.0	1850	6	ABQ74193	Zebrafish	c 208	13.8	76.7	50	12	ADM32608	Adm32608	HPIV2 T88
136	14.4	80.0	1850	6	ABQ74193	Zebrafish	c 209	13.8	76.7	51	4	AAH90180	Aah90180	HPIV2 Gre
137	14.4	80.0	1850	6	ABQ74193	Zebrafish	c 210	13.8	76.7	60	6	ABN44820	Abn44820	Human cto
138	14.4	80.0	1883	11	AAK51626	Human pol	c 211	13.8	76.7	103	3	ABN44820	Abn44820	Human cto
139	14.4	80.0	1883	11	ADPE5676	Human fib	c 212	13.8	76.7	103	3	ABN44820	Abn44820	Human cto
140	14.4	80.0	1885	11	ADPE5552	Human fib	c 213	13.8	76.7	200	4	ABA70112	Aba70112	Tomato pl
141	14.4	80.0	1885	11	AAAS40978	Human fib	c 214	13.8	76.7	200	4	ABA70112	Aba70112	Human foe
142	14.4	80.0	1918	11	ADPE5102	Human fib	c 215	13.8	76.7	200	4	AAI50231	Aai50231	Probe #18
143	14.4	80.0	1918	11	ADPE5102	Human fib	c 216	13.8	76.7	200	4	ABA56898	Abas6898	Probe #15
144	14.4	80.0	1935	4	ABL27239	Drosophi1	c 217	13.8	76.7	200	4	AAK44226	Aak44226	Human bon
145	14.4	80.0	2111	4	AAK51625	Human pol	c 218	13.8	76.7	200	4	AAK18327	Aak18327	Human bra
146	14.4	80.0	2120	5	AAK51624	Human pol	c 219	13.8	76.7	200	6	ABG43883	Abg43883	Human liv
147	14.4	80.0	2484	5	AAK51624	Human pol	c 220	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
148	14.4	80.0	2553	4	AAK52608	Human pol	c 221	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
149	14.4	80.0	2553	4	AAK52609	Human pol	c 222	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
150	14.4	80.0	2553	4	AAK52610	Human pol	c 223	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
151	14.4	80.0	2553	4	AAK52610	Human pol	c 224	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
152	14.4	80.0	2553	4	AAK52610	Human pol	c 225	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
153	14.4	80.0	2553	4	AAK52610	Human pol	c 226	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
154	14.4	80.0	2553	4	AAK52610	Human pol	c 227	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
155	14.4	80.0	2553	4	AAK52610	Human pol	c 228	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
156	14.4	80.0	2553	4	AAK52610	Human pol	c 229	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
157	14.4	80.0	2553	4	AAK52610	Human pol	c 230	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
158	14.4	80.0	2553	4	AAK52610	Human pol	c 231	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
159	14.4	80.0	2553	4	AAK52610	Human pol	c 232	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
160	14.4	80.0	2553	4	AAK52610	Human pol	c 233	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
161	14.4	80.0	2553	4	AAK52610	Human pol	c 234	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
162	14.4	80.0	2553	4	AAK52610	Human pol	c 235	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
163	14.4	80.0	2553	4	AAK52610	Human pol	c 236	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
164	14.4	80.0	2553	4	AAK52610	Human pol	c 237	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
165	14.4	80.0	2553	4	AAK52610	Human pol	c 238	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
166	14.4	80.0	2553	4	AAK52610	Human pol	c 239	13.8	76.7	202	2	AAH85659	Aah85659	Human gin
167	14.4	80.0	2553	4	AAK52610	Human pol	c 240	13.8	76.7	202	2	AAH85659	Aah85659	Human gin

C 241	13.8	76.7	474	4	ABA57842	AbA57842 Human foe	C 314	13.8	76.7	1041	3	AAZ99484	AAZ99484 cDNA enco
C 242	13.8	76.7	474	4	AAI37038	AAI37038 Probe #57	C 315	13.8	76.7	1041	6	AAD40274	AAD40274 Pumpkin 2
C 243	13.8	76.7	474	4	AAI37433	AAI37433 Probe #61	C 316	13.8	76.7	1046	3	AAC99067	AAC99067 Human pan
C 244	13.8	76.7	474	4	ABA27182	ABA27182 Probe #56	C 317	13.8	76.7	1053	6	ABN98389	ABN98389 Arabidops
C 245	13.8	76.7	474	4	AAK31131	AAK31131 Human bon	C 318	13.8	76.7	1086	3	AAC38741	AAC38741 Arabidops
C 246	13.8	76.7	474	4	AAK31554	AAK31554 Human bon	C 319	13.8	76.7	1146	6	AAD23611	AAD23611 Canine p5
C 247	13.8	76.7	474	4	AAK05916	AAK05916 Human bra	C 320	13.8	76.7	1166	2	AAZ42037	AAZ42037 Human end
C 248	13.8	76.7	474	4	AAK05529	AAK05529 Human bra	C 321	13.8	76.7	1174	6	AAD23609	AAD23609 Canine p5
C 249	13.8	76.7	474	4	ABN30811	ABN30811 Human liv	C 322	13.8	76.7	1207	3	AAC45575	AAC45575 Arabidops
C 250	13.8	76.7	474	4	ABN31238	ABN31238 Human liv	C 323	13.8	76.7	1210	3	AAC35305	AAC35305 Arabidops
C 251	13.8	76.7	474	6	ABN05883	ABN05883 Human gen	C 324	13.8	76.7	1210	8	ACA53938	ACA53938 Prokaryot
C 252	13.8	76.7	474	6	ABN06308	ABN06308 Human gen	C 325	13.8	76.7	1272	8	ACA47824	ACA47824 Prokaryot
C 253	13.8	76.7	475	5	ABV37833	ABV37833 Human pro	C 326	13.8	76.7	1296	3	AAC51827	AAC51827 Arabidops
C 254	13.8	76.7	486	4	AAO07135	AAO07135 DNA enco	C 327	13.8	76.7	1305	8	ACA32204	ACA32204 Prokaryot
C 255	13.8	76.7	487	4	AAO41066	AAO41066 Zea may	C 328	13.8	76.7	1332	3	AAC42359	AAC42359 Arabidops
C 256	13.8	76.7	491	9	ACH33647	ACH33647 Human end	C 329	13.8	76.7	1373	12	ADH70091	ADH70091 Human Vbe
C 257	13.8	76.7	494	10	ADF85661	ADF85661 Human ade	C 330	13.8	76.7	1395	6	ABZ12919	ABZ12919 Arabidops
C 258	13.8	76.7	501	4	ABA60737	ABA60737 Human probe	C 331	13.8	76.7	1395	12	ADN73174	ADN73174 Thale cre
C 259	13.8	76.7	501	4	AAI40628	AAI40628 Probe #93	C 332	13.8	76.7	1431	8	ACC51103	ACC51103 Human Dop
C 260	13.8	76.7	501	4	AAK34912	AAK34912 Human bon	C 333	13.8	76.7	1431	8	ACC51104	ACC51104 Human Dop
C 261	13.8	76.7	501	4	AAK09021	AAK09021 Human bra	C 334	13.8	76.7	1457	10	ADC32127	ADC32127 Human nov
C 262	13.8	76.7	501	4	ABN34668	ABN34668 Human liv	C 335	13.8	76.7	1463	3	AAC52279	AAC52279 Arabidops
C 263	13.8	76.7	514	4	AAFI17929	AAFI17929 Human bre	C 336	13.8	76.7	1493	3	AAC77153	AAC77153 Human ORF
C 264	13.8	76.7	514	4	AAI47359	AAI47359 Human bre	C 337	13.8	76.7	1506	6	ABA91373	ABA91373 Human bre
C 265	13.8	76.7	514	6	ABN63960	ABN63960 Human bre	C 338	13.8	76.7	1513	6	ABK46131	ABK46131 cDNA enco
C 266	13.8	76.7	514	10	ABT33172	ABT33172 Human tum	C 339	13.8	76.7	1513	8	ACC50153	ACC50153 Breast ca
C 267	13.8	76.7	514	10	ADL33079	ADL33079 Human bre	C 340	13.8	76.7	1513	10	ADB75292	ADB75292 Prostate
C 268	13.8	76.7	514	12	ADL44369	ADL44369 Human cDN	C 341	13.8	76.7	1518	3	AAZ49454	AAZ49454 Mouse Vol
C 269	13.8	76.7	516	11	ACH96986	ACH96986 Klebsell	C 342	13.8	76.7	1545	12	ADP04788	ADP04788 Sea equir
C 270	13.8	76.7	519	4	AAFI17770	AAFI17770 Human bre	C 343	13.8	76.7	1548	3	AAAC48768	AAAC48768 Arabidops
C 271	13.8	76.7	519	6	ABT07200	ABT07200 Human bre	C 344	13.8	76.7	1551	3	AAC32972	AAC32972 Arabidops
C 272	13.8	76.7	519	4	ABT08855	ABT08855 Human bre	C 345	13.8	76.7	1552	3	AAZ50890	AAZ50890 Human rec
C 273	13.8	76.7	519	6	ABN63801	ABN63801 Human bre	C 346	13.8	76.7	1626	5	AAI56192	AAI56192 DNA enco
C 274	13.8	76.7	519	11	ADL33013	ADL33013 Human tum	C 347	13.8	76.7	1626	5	AAI57714	AAI57714 DNA enco
C 275	13.8	76.7	519	12	ADL44210	ADL44210 Human bre	C 348	13.8	76.7	1677	8	ABX63140	ABX63140 Human cDN
C 276	13.8	76.7	519	12	AAI576168	AAI576168 DNA enco	C 349	13.8	76.7	1715	10	ADFI17114	ADFI17114 Lettuce c
C 277	13.8	76.7	522	5	AAI576168	AAI576168 DNA enco	C 350	13.8	76.7	1722	2	AAT004002	AAT004002 Squale
C 278	13.8	76.7	529	5	ABV15936	ABV15936 Human pro	C 351	13.8	76.7	1724	12	ADH13717	ADH13717 Human ENZ
C 279	13.8	76.7	544	10	ACD93052	ACD93052 Human col	C 352	13.8	76.7	1734	2	AAQ97800	AAQ97800 Clone pCB
C 280	13.8	76.7	565	5	ABV45736	ABV45736 Human pro	C 353	13.8	76.7	1734	6	ABT99373	ABT99373 Mouse lsc
C 281	13.8	76.7	589	5	ABV56962	ABV56962 Human pro	C 354	13.8	76.7	1741	3	AAC32892	AAC32892 Arabidops
C 282	13.8	76.7	598	3	AAI32814	AAI32814 Arabidops	C 355	13.8	76.7	1745	3	AAAC45529	AAAC45529 Arabidops
C 283	13.8	76.7	620	10	ADD33713	ADD33713 Mouse mit	C 356	13.8	76.7	1794	10	ADC30223	ADC30223 Human nov
C 284	13.8	76.7	640	10	ADG37560	ADG37560 Aspergill	C 357	13.8	76.7	1831	4	ABT04375	ABT04375 Drosophi
C 285	13.8	76.7	649	4	AAH72227	AAH72227 Human cer	C 358	13.8	76.7	1841	12	ADH13220	ADH13220 Human mal
C 286	13.8	76.7	687	5	ABV28360	ABV28360 Human pro	C 359	13.8	76.7	1853	2	AAV73011	AAV73011 Human adu
C 287	13.8	76.7	687	5	ABV22545	ABV22545 Human pro	C 360	13.8	76.7	1853	6	ABO92036	ABO92036 Human pol
C 288	13.8	76.7	690	10	ADD33716	ADD33716 Mouse mit	C 361	13.8	76.7	1893	4	AAH16343	AAH16343 Human cDN
C 289	13.8	76.7	700	6	ABK46132	ABK46132 cDNA enco	C 362	13.8	76.7	1938	2	AAZ42120	AAZ42120 Human end
C 290	13.8	76.7	758	4	AAH05160	AAH05160 Human cDN	C 363	13.8	76.7	1956	2	AAQ88712	AAQ88712 Aspergill
C 291	13.8	76.7	765	3	AAZ47241	AAZ47241 HTLV-1 Ta	C 364	13.8	76.7	2000	10	ADC08567	ADC08567 Rice DNA
C 292	13.8	76.7	765	3	AAZ47276	AAZ47276 HLA-A2/HT	C 365	13.8	76.7	2002	3	AACT7836	AACT7836 Human can
C 293	13.8	76.7	765	3	AAZ56677	AAZ56677 TCR alpha	C 366	13.8	76.7	2020	2	AAAX00649	AAAX00649 Human sec
C 294	13.8	76.7	765	3	AAZ56667	AAZ56667 TCR alpha	C 367	13.8	76.7	2067	8	ACA39054	ACA39054 Prokaryot
C 295	13.8	76.7	778	4	AAH71164	AAH71164 Human cer	C 368	13.8	76.7	2070	12	ADL04080	ADL04080 DNA enco
C 296	13.8	76.7	792	4	ABZ14145	ABZ14145 Arabidops	C 369	13.8	76.7	2137	4	ABL15573	ABL15573 Drosophi
C 297	13.8	76.7	796	2	AAI29757	AAI29757 D10 singl	C 370	13.8	76.7	2163	4	ABN66445	ABN66445 Streptoco
C 298	13.8	76.7	859	4	AAI35629	AAI35629 Human mus	C 371	13.8	76.7	2179	6	ABQ93371	ABQ93371 Human cod
C 299	13.8	76.7	859	4	ABX58617	ABX58617 cDNA enco	C 372	13.8	76.7	2199	10	ADBS8258	ADBS8258 Toxicity-
C 300	13.8	76.7	859	12	ADJ28344	ADJ28344 Human mus	C 373	13.8	76.7	2199	10	ADBS52789	ADBS52789 Primary r
C 301	13.8	76.7	870	5	AAI572842	AAI572842 DNA enco	C 374	13.8	76.7	2199	10	ADBS5175	ADBS5175 Rat equal
C 302	13.8	76.7	870	5	AAI572842	AAI572842 DNA enco	C 375	13.8	76.7	2225	12	ADP72506	ADP72506 Renal tox
C 303	13.8	76.7	905	6	ABN65888	ABN65888 Lung canc	C 376	13.8	76.7	2225	4	ABL03473	ABL03473 Drosophi
C 304	13.8	76.7	905	6	ABN93926	ABN93926 Gene #424	C 377	13.8	76.7	2238	12	ADJ39529	ADJ39529 Plant cDN
C 305	13.8	76.7	936	5	AAI570137	AAI570137 DNA enco	C 378	13.8	76.7	2307	8	ABX71087	ABX71087 Novel hum
C 306	13.8	76.7	949	3	AAI570137	AAI570137 DNA enco	C 379	13.8	76.7	2420	5	ABV25077	ABV25077 Human pro
C 307	13.8	76.7	949	3	AAI570137	AAI570137 DNA enco	C 380	13.8	76.7	2420	5	ABV25077	ABV25077 Human pro
C 308	13.8	76.7	949	10	ABZ96953	ABZ96953 Human low	C 381	13.8	76.7	2465	4	AAH16721	AAH16721 Human cDN
C 309	13.8	76.7	949	11	ABD20802	ABD20802 Human pul	C 382	13.8	76.7	2510	11	ADM02669	ADM02669 Human cDN
C 310	13.8	76.7	964	4	AAH02697	AAH02697 Human col	C 383	13.8	76.7	2513	3	AAH14259	AAH14259 Human cDN
C 311	13.8	76.7	967	4	AAH33694	AAH33694 Human col	C 384	13.8	76.7	2568	8	ABT18130	ABT18130 Aspergill
C 312	13.8	76.7	1016	6	ABK34682	ABK34682 Human cDN	C 385	13.8	76.7	2683	8	ABT19944	ABT19944 Aspergill
C 313	13.8	76.7	1032	6	ABN67322	ABN67322 Breast ep	C 386	13.8	76.7	2718	4	AAI37575	AAI37575 Human mus

387	13.8	76.7	2718	8	ABX60563	Abx60563 cDNA enco	460	13.8	76.7	7473	12	ADQ33420	Adq33420 Human sof
388	13.8	76.7	2718	12	ADJ31313	Adj31313 Human mus	461	13.8	76.7	9596	4	AAK66614	Aak66614 Human imm
389	13.8	76.7	2792	4	ABL05249	Abi05249 Drosophi	462	13.8	76.7	9596	4	AAK66613	Aak66613 Human imm
390	13.8	76.7	2825	6	AAD29162	Aad29162 Sheep K2.	463	13.8	76.7	9903	4	AAL05500	Aal05500 Human rep
391	13.8	76.7	2825	8	ACC49767	Acc49767 Ovine K2.	464	13.8	76.7	10785	4	ABA07367	Abao7367 Human pan
392	13.8	76.7	2831	4	ABL06391	Abi06391 Drosophi	465	13.8	76.7	10785	4	AAK91098	Aak91098 Human pan
393	13.8	76.7	2859	12	ADP99210	Adp99210 Human tra	466	13.8	76.7	10785	4	AAK54411	Aak54411 Human dig
394	13.8	76.7	3019	5	AA583874	Aa583874 DNA enco	467	13.8	76.7	10785	5	AA530060	Aa530060 Human imm
395	13.8	76.7	3104	4	AAH15749	Aah15749 Human cdn	468	13.8	76.7	10785	10	ADB33397	Adb33397 Human nov
396	13.8	76.7	3114	4	AAH15789	Aah15789 Human cdn	469	13.8	76.7	10785	11	ADN97366	Adn97366 B. lichen
397	13.8	76.7	3146	1	AAH81713	Aah81713 Clone con	470	13.8	76.7	13315	4	ABA07371	Abao7371 Human pan
398	13.8	76.7	3166	11	ADM02054	Adm02054 Human cdn	471	13.8	76.7	13315	4	AAK91102	Aak91102 Human dig
399	13.8	76.7	3194	12	ADQ24917	Adq24917 Human sof	472	13.8	76.7	13315	4	AAK54415	Aak54415 Human imm
400	13.8	76.7	3267	12	ADN72212	Adn72212 Thale cre	473	13.8	76.7	13315	5	AA530064	Aa530064 Human lun
401	13.8	76.7	3349	8	ABZ73720	Abz73720 Secreted	474	13.8	76.7	13315	10	ADB33401	Adb33401 Human nov
402	13.8	76.7	3349	10	ADC20594	Adc20594 Human sec	475	13.8	76.7	14141	4	ABL06822	Abi06822 Drosophi
403	13.8	76.7	3349	10	ABT16868	Abt16868 Human sec	476	13.8	76.7	14141	12	ADM32597	Adm32597 Human par
404	13.8	76.7	3349	10	ABT16868	Abt16868 Human sec	477	13.8	76.7	15654	12	ADM32595	Adm32595 Human par
405	13.8	76.7	3350	8	ABZ73719	Abz73719 Secreted	478	13.8	76.7	15654	12	ADM32596	Adm32596 Human par
406	13.8	76.7	3350	10	ABT16867	Abt16867 Human sec	479	13.8	76.7	16084	4	AAK69442	Aak69442 Human imm
407	13.8	76.7	3350	10	ABT16867	Abt16867 Human sec	480	13.8	76.7	16084	4	AAK82881	Aak82881 Human imm
408	13.8	76.7	3350	10	ABT16867	Abt16867 Human sec	481	13.8	76.7	16084	4	AAK85311	Aak85311 Human imm
409	13.8	76.7	3443	12	ADN04827	Adn04827 Act1pso	482	13.8	76.7	19497	4	ABA07368	Abao7368 Human pan
410	13.8	76.7	3734	3	AAA35138	Aaa35138 Human ade	483	13.8	76.7	20261	4	AAK91099	Aak91099 Human dig
411	13.8	76.7	3734	6	AAK83777	Aak83777 Human low	484	13.8	76.7	20261	4	AAK65412	Aak65412 Human imm
412	13.8	76.7	3734	6	ABK83777	Abk83777 Human cdn	485	13.8	76.7	20261	5	AA530061	Aa530061 Human lun
413	13.8	76.7	3734	11	ABZ96954	Abz96954 Human nuc	486	13.8	76.7	20261	10	ADB33398	Adb33398 Human nov
414	13.8	76.7	3734	11	ADP17112	Adp17112 Lettuce c	487	13.8	76.7	24187	4	AA595929	Aa595929 Proponib
415	13.8	76.7	3839	10	AAH72798	Aah72798 Human cer	488	13.8	76.7	24187	8	ACF64458	Acf64458 Proponib
416	13.8	76.7	3856	5	ADL65339	Adl65339 Human ova	489	13.8	76.7	24550	4	AA595608	Aa595608 Proponib
417	13.8	76.7	3900	4	ABL04374	Abi04374 Drosophi	490	13.8	76.7	24550	8	ACF64537	Acf64537 Proponib
418	13.8	76.7	4345	4	AAK75170	Aak75170 Human imm	491	13.8	76.7	32223	5	ABA19863	Abai19863 Human ner
419	13.8	76.7	4345	4	AAK75169	Aak75169 Human imm	492	13.8	76.7	32223	5	ABA19863	Abai19863 Human ner
420	13.8	76.7	4357	2	AAV23493	Aav23493 Pseudomon	493	13.8	76.7	32835	9	ADB72686	Adb72686 Human NEK
421	13.8	76.7	4377	3	AAA13904	Aaa13904 Pseudomon	494	13.8	76.7	32835	10	ADB72686	Adb72686 Human NEK
422	13.8	76.7	4377	3	AAAF30869	Aaf30869 Pseudomon	495	13.8	76.7	32835	12	ADM74543	Adm74543 Human nek
423	13.8	76.7	4377	6	AAD22881	Aad22881 P. alcali	496	13.8	76.7	32835	12	ADM74543	Adm74543 Human nek
424	13.8	76.7	4377	6	AAD22881	Aad22881 P. alcali	497	13.8	76.7	51001	12	ADJ10262	Adj10262 Human ger
425	13.8	76.7	4518	9	ABT43709	Abt43709 Molecule	498	13.8	76.7	55996	9	ADA02741	Ada02741 Mouse Tnf
426	13.8	76.7	4539	4	ABL03472	Abi03472 Drosophi	499	13.8	76.7	55996	10	ADB72479	Adb72479 Mouse Tnf
427	13.8	76.7	4667	4	ABL05017	Abi05017 Drosophi	500	13.8	76.7	55996	10	ADC85221	Adc85221 Mouse Tnf
428	13.8	76.7	4683	3	AAH35139	Aah35139 Human ade							
429	13.8	76.7	4683	3	AAH35139	Aah35139 Human ade							
430	13.8	76.7	4683	10	ABZ96955	Abz96955 Human low							
431	13.8	76.7	4683	3	AAH35139	Aah35139 Human ade							
432	13.8	76.7	4683	11	ABD20804	Abd20804 Human pul							
433	13.8	76.7	4683	11	ABD20804	Abd20804 Human pul							
434	13.8	76.7	4920	4	ABL15572	Abi15572 Drosophi							
435	13.8	76.7	4928	4	ABL06823	Abi06823 Drosophi							
436	13.8	76.7	5240	2	AAV48268	Aav48268 P. chryso							
437	13.8	76.7	5377	4	ABL16036	Abi16036 Drosophi							
438	13.8	76.7	5439	4	ABL05248	Abi05248 Drosophi							
439	13.8	76.7	5439	5	ABA19882	Abai19882 Human ner							
440	13.8	76.7	5825	5	ACH03918	Ach03918 Human cdn							
441	13.8	76.7	6102	10	ADBS5284	Adbs5284 Rat gene							
442	13.8	76.7	6102	10	ADBS5300	Adbs5300 Rat gene							
443	13.8	76.7	6102	10	ADBS5292	Adbs5292 Rat gene							
444	13.8	76.7	6102	10	ADBS5304	Adbs5304 Rat gene							
445	13.8	76.7	6102	10	ADBS5286	Adbs5286 Rat gene							
446	13.8	76.7	6393	2	AAZ23079	Aaz23079 Human pta							
447	13.8	76.7	6393	2	AAZ23079	Aaz23079 Human pta							
448	13.8	76.7	6393	8	ACC50261	Acc50261 Breast ca							
449	13.8	76.7	6438	8	ABZ73721	Abz73721 Secreted							
450	13.8	76.7	6438	10	ADC20595	Adc20595 Human sec							
451	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
452	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
453	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
454	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
455	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
456	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
457	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
458	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							
459	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec							

## ALIGNMENTS

RESULT 1	
ID	AAQ88280 standard; cDNA; 18 BP.
AC	AAQ88280;
XX	
DT	25-MAR-2003 (revised)
DT	06-DEC-1995 (first entry)
XX	
DE	Antisense oligonucleotide to inhibit p53 gene expression.
XX	
KW	Antisense; recombinant virus; vector; p53; binding site; Huntington's;
KW	neuronal degeneration; Parkinson's; Alzheimer's;
KW	amyotrophic lateral sclerosis; ss.
XX	
OS	Synthetic.
XX	
PN	MO950916-A1.
XX	
PD	13-APR-1995.
XX	
PF	29-SEP-1994; 94WO-FR001142.
XX	
PR	04-OCT-1993; 93FR-00011774.
XX	
PA	(RHON ) RHONE POULENC RORER SA.
XX	



PI Mallet J, Revah F, Stutzmann J;  
XX  
XX WPI; 1995-155257/20.  
XX  
XX Treating and preventing neurodegenerative diseases - with cpds. that  
PT inhibit p53 activity, partic. an anti-sense sequence or viral vector.  
XX  
XX Example 1; Page 10; 21pp; French.  
XX  
XX The sequence of an antisense oligonucleotide used either solely or as  
CC part of a recombinant viral vector to inhibit the expression of the p53  
CC gene. The oligonucleotide may be used in conjunction either as naked DNA  
CC or on the same recombinant viral vector, with the p53 binding sequence  
CC (AAQ88279). The viral vector may also encode a mutated form of p53 which  
CC antagonises p53 activity. These oligonucleotides and recombinant viruses  
CC can be used in compositions to treat and prevent neuronal degeneration,  
CC Huntington's, Parkinson's or Alzheimer's diseases or amyotrophic lateral  
CC sclerosis. (updated on 25-MAR-2003 to correct FN field.)  
XX  
XX Sequence 18 BP; 4 A; 6 C; 3 G; 5 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CGACTGTGATCTCTCCAT 18  
DB 1 CGACTGTGATCTCTCCAT 18  
RESULT 2  
AAZ52303/c  
ID AAZ52303 standard; cDNA; 1416 BP.  
XX  
XX AAZ52303;  
AC  
XX 24-JUL-2000 (first entry)  
DT  
XX  
XX Rat p53 (p53 isoform) cDNA.  
DE  
XX  
XX Rat p53 isoform; p53; reverse transcription-PCR; marker; hypoxia;  
KM myocardial infarction; cell proliferation; cytostatic; proliferative; ss.  
XX  
XX Rattus norvegicus.  
OS  
XX  
XX Key Location/Qualifiers  
FH 24.838  
FT /tag= a  
FT /product= "Rat p53 protein"  
FT /transl\_except= (pos: 759.760, aa: Arg)  
FT /transl\_except= (pos: 821.823, aa: Ala)  
XX  
XX MO200022127-A1.  
PN  
XX  
XX 20-APR-2000.  
PD  
XX  
XX 07-OCT-1999; 99WO-US023319.  
PE  
XX  
XX 09-OCT-1998; 98US-0103849P.  
PR  
XX  
XX (BGMH) BRIGHAM & WOMENS HOSPITAL INC.  
PA  
XX  
XX Dell'acqua G, Mann MJ, Dzeu VJ;  
PI  
XX  
XX WPI; 2000-317984/27.  
DR  
XX  
XX P-PSDB; AAV70713.  
PT  
XX  
XX Novel isoform of p53 useful as a marker of myocardial infarction and for  
PT controlling cellular proliferation.  
XX  
XX Claim 7; Fig 2; 25pp; English.  
PS  
XX  
XX The present cDNA sequence encodes the rat p53 isoform, p53 derived from

CC rat cardiac myocytes by reverse transcription-PCR of RNA. This p53  
CC isoform is truncated to eliminate a substantial portion of the C-terminal  
CC end of the protein. p53 is useful as a marker of myocardial infarction  
CC (by the indication of hypoxia) and in the control of cell proliferation  
CC both in vivo and in vitro. It was able to increase the transactivation  
CC achieved through co-transfection with wild type p53. The combination of  
CC p53 with wild type p53 produce an enhanced inhibition of primary  
CC incorporation at 24 hours compared to wild type p53 alone  
XX  
XX Sequence 1416 BP; 312 A; 432 C; 323 G; 349 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 18; DB 3; Length 1416;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CGACTGTGATCTCTCCAT 18  
DB 41 CGACTGTGATCTCTCCAT 24  
RESULT 3  
ABK63584/c  
ID ABK63584 standard; cDNA; 1627 BP.  
XX  
XX  
XX ABK63584;  
AC  
XX  
XX 18-JUN-2002 (first entry)  
DT  
XX  
XX Rat sequence differentially expressed in response to a hepatotoxin #1491.  
DE  
XX  
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
KM differential expression; centrilobular necrosis; steatosis.  
XX  
XX Rattus norvegicus.  
OS  
XX  
XX MO200210453-A2.  
PN  
XX  
XX 07-FEB-2002.  
PD  
XX  
XX 30-JUL-2001; 2001WO-US023872.  
PF  
XX  
XX 31-JUL-2000; 2000US-0222040P.  
PR  
XX  
XX 02-NOV-2000; 2000US-0244880P.  
PR  
XX  
XX 11-MAY-2001; 2001US-0230029P.  
PR  
XX  
XX 15-MAY-2001; 2001US-0230645P.  
PR  
XX  
XX 22-MAY-2001; 2001US-0232336P.  
PR  
XX  
XX 06-JUN-2001; 2001US-0235738P.  
PR  
XX  
XX 13-JUN-2001; 2001US-0237457P.  
PR  
XX  
XX 19-JUN-2001; 2001US-0238884P.  
PR  
XX  
XX 09-JUL-2001; 2001US-0303459P.  
XX  
XX (GENE-) GENE LOGIC INC.  
PA  
XX  
XX  
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;  
PI  
XX  
XX WPI; 2002-241625/29.  
DR  
XX  
XX  
XX Predicting toxic effects of compounds or the progression of these toxic  
PT effects by determining the changes in gene expression in tissues or cells  
PT exposed to the toxin and comparing these to gene expression in unexposed  
PT tissues or cells.  
XX  
XX  
XX Claim 1; SEQ ID NO 1491; 239pp; English.  
PS  
XX  
XX The invention relates to methods for predicting toxic effects of  
CC compounds or the progression of these toxic effects by determining the  
CC global changes in gene expression in tissues or cells exposed to the  
CC toxin and comparing these to gene expression in unexposed tissues or  
CC cells. Also included are methods of predicting at least one toxic effect  
CC of a compound or progression of a toxic effect, preferably the  
CC hepatotoxicity of a compound, comprising detecting the level of  
CC expression in a tissue or cell sample exposed to the compound of two or  
CC more genes listed in the specification, where differential expression of

CC the genes is indicative of at least one toxic effect or progression. The  
CC method can also be used to identify an agent which modulates the toxic  
CC response and predict cellular pathways that a compound modulates in a  
CC cell. The methods utilize a set of at least two probes (on a solid  
CC support in kit form), where each of the probes comprises a sequence that  
CC specifically hybridises to a gene listed in the specification, a computer  
CC system comprising a database containing information identifying the  
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
CC set of genes comprising at least two genes listed in the specification,  
CC and a user interface to view the information used to present information  
CC identifying the expression level in a tissue or cell of at least one gene  
CC listed in the specification. The method is useful for elucidating global  
CC changes in gene expression and for identifying toxicity markers in  
CC tissues or cell exposed to a known toxin. The genes may be used as  
CC toxicity markers in drug screening and toxicity assays. The genes and  
CC gene expression information may be used as diagnostic markers for the  
CC prediction or identification of the physiological state of tissue or cell  
CC sample that has been exposed to a compound or agent. Hepatotoxicity is  
CC characterised by centrilobular necrosis and steatosis. The present  
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
CC which is differentially expressed in response to a hepatotoxic agent

SQ Sequence 1627 BP; 374 A; 486 C; 380 G; 387 T; 0 U; 0 Other;  
Query Match 100.0%; Score 18; DB 6; Length 1627;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCAT 18  
Db 41 CGACTGTGAATCCTCCAT 24

RESULT 4

ADP72925/C  
ID ADP72925 standard; DNA; 1627 BP.

AC ADP72925;

DT 26-AUG-2004 (first entry)

DE Renal toxin progression gene marker #1514.

XX ds; toxic effect; gene expression profile; kidney tissue;

KM differential gene expression; toxicity progression; toxicity marker;

KM drug screening; toxicity assay; kidney pathology; nephritis;

KM kidney necrosis; glomerular injury; tubular injury;

KM focal segmental glomerulosclerosis.

OS Rattus norvegicus.

PN WO2004048598-A2.

XX 10-JUN-2004.

XX 24-NOV-2003; 2003WO-US037556.

XX 22-NOV-2002; 2002US-00301856.

XX (GENE-) GENE LOGIC INC.

XX Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;

XX Elashoff M;

XX WPI; 2004-460771/43.

XX Predicting (the progression of) a toxic effect of a compound, for  
XX monitoring the progression of renal disease states, comprises preparing a  
XX gene expression profile of a kidney tissue or cell sample exposed to the  
XX compound.

XX Claim 11; SEQ ID NO 1514; 266bp; English.

CC The invention relates to a method of predicting (the progression of) a  
CC toxic effect of a compound by preparing a gene expression profile of a  
CC kidney tissue or cell sample exposed to the compound and comparing the  
CC gene expression profile to a database, or detecting the level of gene(s)  
CC expression in a tissue or cell sample exposed to the compound, where  
CC differential gene expression compared to a control indicates a toxic  
CC effect (toxicity progression). The method is useful for predicting (the  
CC progression of) at least one toxic effect of a compound. The genes are  
CC useful as toxicity markers in drug screening and toxicity assays. The  
CC methods are useful for predicting the likelihood that a compound or test  
CC agent will induce various specific kidney pathologies, such as nephritis,  
CC kidney necrosis, glomerular and tubular injury, or focal segmental  
CC glomerulosclerosis. The methods are useful for determining the similarity  
CC of a toxic response to one or more individual compounds and for  
CC predicting or elucidating the potential cellular pathways influenced,  
CC induced or modulated by the compound or test agent. The kit is useful for  
CC predicting or modelling the toxic response of a test compound, for  
CC monitoring the progression of renal disease states, for identifying genes  
CC that show promise as new drug targets and for screening known and newly  
CC designed drugs. This sequence corresponds to a gene marker used in the  
CC method of the invention. (Note: The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences).

SQ Sequence 1627 BP; 374 A; 486 C; 380 G; 387 T; 0 U; 0 Other;  
Query Match 100.0%; Score 18; DB 12; Length 1627;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCAT 18  
Db 41 CGACTGTGAATCCTCCAT 24

RESULT 5

ADB58550/C  
ID ADB58550 standard; DNA; 1786 BP.

AC ADB58550;

DT 04-DEC-2003 (first entry)

DE Toxicity-related gene, SEQ ID 3576.

XX Toxic; toxin; gene expression profile; hepatotoxicity; liver;

KM drug screening; toxicity assay; ds.

XX Unidentified.

OS WO2003064624-A2.

PN 07-AUG-2003.

XX 31-JAN-2003; 2003WO-US003194.

XX 31-JAN-2002; 2002US-00060087.

XX 15-MAR-2002; 2002US-0364045P.

XX 15-MAR-2002; 2002US-0364055P.

XX 30-DEC-2002; 2002US-0436643P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;

XX WPI; 2003-689530/65.

XX Predicting a toxic effect of a compound, useful in identifying toxicity  
XX markers in liver tissues or cells for drug screening and toxicity assays,  
XX comprises preparing gene expression profile of tissue or cells exposed to  
XX the compound.

XX Claim 1; SEQ ID NO 3576; 1156bp; English.

XX The present invention relates to a method for predicting a toxic effect  
 CC of a compound. The method comprises preparing a gene expression profile  
 CC of a tissue or cell sample exposed to the compound, and comparing the  
 CC gene expression profile to a database comprising SEQ ID 1-4925, where  
 CC differential expression of the gene indicates at least one toxic effect.  
 CC The method is useful for predicting at least one toxic effect of a  
 CC compound, predicting hepatotoxicity or the progression of a toxic effect  
 CC of a compound, identifying an agent that modulates the onset or  
 CC progression of a toxic response, predicting the cellular pathways that a  
 CC compound modulates in a cell, and identifying an agent that modulates at  
 CC least one activity of a protein. The method and compositions of the  
 CC present invention using a database of genes having liver toxin-induced  
 CC differential expression, are useful in identifying toxicity markers in  
 CC liver tissues or cells for drug screening and toxicity assays. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 1786 BP; 400 A; 533 C; 425 G; 428 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 10; Length 1786;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18

DB 201 CGACTGTGAATCCTCCAT 184

RESULT 6

ADBS3157/C ADBS3157 standard; DNA; 1786 BP.

XX ADBS3157;

DT 04-DEC-2003 (first entry)

XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3699.

XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;

KW toxicity marker; toxicity progression; drug screening;

KW primary rat hepatocyte toxicity modelling; gene; de.

OS Rattus norvegicus.

XX WO2003065993-A2.

XX 14-AUG-2003.

PD 04-FEB-2003; 2003WO-US003482.

XX 04-FEB-2002; 2002US-0353171P.

PR 13-MAR-2002; 2002US-036534P.

PR 08-APR-2002; 2002US-0370248P.

PR 10-APR-2002; 2002US-0371134P.

PR 10-APR-2002; 2002US-0371135P.

PR 10-APR-2002; 2002US-0371150P.

PR 11-APR-2002; 2002US-0371413P.

PR 19-APR-2002; 2002US-0373601P.

PR 19-APR-2002; 2002US-0374139P.

PR 08-MAY-2002; 2002US-0378370P.

PR 09-MAY-2002; 2002US-0378652P.

PR 09-MAY-2002; 2002US-0378653P.

PR 09-JUL-2002; 2002US-0394230P.

PR 09-JUL-2002; 2002US-0394253P.

PR 04-SEP-2002; 2002US-0407688P.

PR 28-JAN-2003; 2003US-0442900P.

XX (GENE-) GENE LOGIC INC.

PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;  
 PI Elashoff M;  
 DR WPI; 2003-731472/69.

XX Determining if a compound induces a toxic effect on a tissue or cell, for  
 PT identifying hepatotoxic compounds, comprises comparing a gene expression  
 PT profile of a tissue or cell sample to a database of Tox mean and non-Tox  
 PT mean values.

XX Claim 44; SEQ ID NO 3699; 874bp; English.

XX The present invention describes a method for determining whether a  
 CC compound induces a toxic effect on a tissue or cell. The method comprises  
 CC preparing a gene expression profile of a tissue or cell sample exposed to  
 CC the compound, and comparing the gene expression profile to a database  
 CC comprising data or information on the Tox mean and non-Tox mean value.  
 CC The method is useful for predicting or identifying at least one toxic  
 CC effect, particularly hepatotoxicity, of a test or unknown compound. The  
 CC genes listed in the specification are useful as diagnostic or toxicity  
 CC markers for the prediction or identification of the physiological state  
 CC of tissue or cell sample that has been exposed to a compound, or to  
 CC identify or predict the toxic effects of a compound or an agent. These  
 CC may also be used as markers for monitoring toxicity progression or for  
 CC drug screening. The present sequence represents a primary rat hepatocyte  
 CC toxicity modelling related gene sequence from the present invention.

SQ Sequence 1786 BP; 400 A; 533 C; 425 G; 428 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 10; Length 1786;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18

DB 201 CGACTGTGAATCCTCCAT 184

RESULT 7

ADL71108/C ADL71108 standard; DNA; 93 BP.

XX ADL71108;

DT 01-JUL-2004 (first entry)

XX p53 hairpin oligo p53-2.

XX bs; p53; hairpin; stem cell; short hairpin RNA; shRNA; anti-HIV;

KW cytostatic; cell therapy; HIV; cancer.

OS Unidentified.

XX WO2004029219-A2.

XX 08-APR-2004.

PF 29-SEP-2003; 2003WO-US030901.

PR 27-SEP-2002; 2002US-0414605P.

XX (COLD-) COLD SPRING HARBOR LAB.

XX Fridman J, Hannon GJ, Hemann M, Lowe SW, Paddison PJ, Zilfou J,

PI Dickinson R, Carmell MA, Rosenquist TA;

DR WPI; 2004-316101/29.

PT Introducing into a subject stem cells with partial or complete loss of  
 PT function of a target gene comprises introducing nucleic acids encoding  
 PT short hairpin RNAs into stem cells and introducing the transfected stem  
 PT cells into the subject.

PS Example 1; Page 37; 93pp; English.

CC The invention relates to a novel method for introducing into a subject a  
CC population of stem cells having partial or complete loss of function of a  
CC target gene by introducing a nucleic acid construct encoding a short  
CC hairpin RNA (shRNA) into stem cells to generate transfected stem cells  
CC and introducing the transfected stem cells into the subject. The  
CC transfected stem cells propagate within the subject and retain partial to  
CC complete loss of function of the target gene. The invention has anti-HIV  
CC and cytostatic activity, and may have a use in cell therapy. The  
CC composition may be administered intravenously or intraarterially, or into  
CC the liver (via portal vein injection), the peritoneal cavity, the kidney  
CC capsule or the bone marrow. The composition and methods are useful for  
CC performing RNA interference in stem cells and for treating diseases (e.g.  
CC HIV or cancer) associated with the expression of a target gene in a  
CC population of cells. The present sequence represents a shRNA nucleic acid  
CC of the invention.

SQ Sequence 93 BP; 26 A; 27 C; 19 G; 21 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 12; Length 93;

Best Local Similarity 94.4%; Pred. No. 75;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18  
|||  
Db 66 CGACTGTGAATCCTCCAT 49

RESULT 8

ABK76382/c  
ID ABK76382 standard; DNA; 390 BP.

AC ABK76382;

DT 13-AUG-2002 (first entry)

DE Bacillus licheniformis genomic sequence tag (GST) #3673.

XX Differential gene expression; genomic sequenced tag; GST;

KM altered culture condition; environmental stress;

OS physiological provocation; ds.

OS Bacillus licheniformis.

PN WO200229113-A2.

PD 11-APR-2002.

PF 05-OCT-2001; 2001WO-US031437.

PR 06-OCT-2000; 2000US-00680598.

PR 27-MAR-2001; 2001US-0279528P.

PA (NOVO) NOVOZYMES BIOTECH INC.

PA (NOVO) NOVOZYMES AS.

PI Berka R, Clausen IG;

PI WPI; 2002-416684/44.

PT Monitoring differential expression of several genes in first Bacillus

PT cell relative to expression of same genes in one or more second Bacillus

PT cells, by using substrate containing Bacillus genomic sequenced tag

PT array.

PS Claim 4; SEQ ID NO 3673; 200pp; English.

CC The invention describes a method of monitoring differential expression of  
CC genes in a first Bacillus cell relative to expression of the genes in  
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes  
CC isolated from Bacillus cells to a substrate containing array of Bacillus  
CC genomic sequenced tags (GST), examining the array, and determining

CC relative gene expression by an observed hybridisation reporter signal of  
CC a spot in the array. The method is useful for measuring the expression of  
CC genes in a first Bacillus cell relative to expression of the same genes  
CC in one or more second Bacillus cells. The method is useful for monitoring  
CC global expression of several genes from a Bacillus cell, discovering new  
CC genes, identifying possible functions of unknown open reading frames and  
CC monitoring gene copy number variation and stability. Monitoring changes  
CC in expression of genes may be used to provide a representation of the way  
CC in which Bacillus cells adapt to changes in culture conditions,  
CC environmental stress or other physiological provocation. Extensive follow  
CC up characterisation is unnecessary, when one spot on an array equals one  
CC gene or one open reading frame, since sequence information is available.  
CC This sequence represents a genomic sequence tag (GST) used in the method  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 390 BP; 96 A; 98 C; 96 G; 100 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 6; Length 390;

Best Local Similarity 94.4%; Pred. No. 85;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18  
|||  
Db 120 CGACTGTGAATCCTCCAT 103

RESULT 9

AAQ67883/c  
ID AAQ67883 standard; DNA; 1173 BP.

AC AAQ67883;

DT 25-MAR-2003 (revised)

DT 23-MAR-1995 (first entry)

DE Murine p53 DNA.

XX Polymerase chain reaction; primer; amplify; NVVAC; ALVAC; recombinant;

KM murine; interleukin-2; IL-2; PRW825; pmu-1; PBS-SK; pmu151; TK vector;

KM plasmid; vaccinia; H6 promoter; amplify; primer; antigenic response;

KM polymerase chain reaction; poxvirus; psd542; immunological response;

KM pathogen; human; interferon; IFN; ss.

XX Synthetic.

OS WO9416716-A1.

PN WO9416716-A1.

PD 04-AUG-1994.

PF 21-JAN-1994; 94WO-US000888.

PR 21-JAN-1993; 93US-00007115.

PR 19-JAN-1994; 94US-00184009.

PA (VIRGO) VIROGENETICS CORP.

PA Paoletti E, Tartaglia J, Cox WI;

PI WPI; 1994-263767/32.

PT Attenuated recombinant virus used for cancer therapy - comprises DNA

PT encoding cytokine and/or tumour associated antigen.

PS Example 31; Fig 38; 232pp; English.

CC This sequence represents the wildtype murine p53 gene from the  
CC translation initiation codon to the stop codon. This sequence was used in  
CC the construction of an ALVAC-based recombinant virus containing the wild-  
CC type murine p53 gene. Three PCR fragments were generated. Fragment I was  
CC produced using primers MM080 and MM081 and plasmid PRW825 as a template.  
CC The 228 bp amplified fragment contains the H6 promoter and the 5'-most

Sequence 1173 BP; 279 A; 349 C; 304 G; 241 T; 0 U; 0 Other;

Qy	1	CGACTGTGATCCTCCAT	18
Db	27	CGACTGTGACTCCTCCAT	10

RESULT 10  
AAZ08528/c  
ID AAZ08528 standard; DNA; 1173 BP

DT 19-OCT-1999 (first entry)

**DB Murine p53 gene.**

KM Attenuated recombinant virus; cytokine; tumour associated antigen;  
KM NYVAC recombinant virus; ALVAC recombinant virus; gene therapy; rabies;  
KM cancer; tumour necrosis factor; nuclear phosphoprotein, p53; IL-2; GM-CSF;  
KM interleukin; interferon; IFN-gamma; IL-4; melanoma associated antigen;  
KM carcinoembryonic antigen; immunisation; antigenic; poxvirus; influenza;  
KM immunological response; immunotherapy; vaccine; Newcastle Disease; ss.  
KM

**Mus sp.**

PN US5942235-A.

PD 24-AUG-1999

PF 02-JUN-1995;

PR 24-DEC-1981

PR 19-JUN-1984;

PR 28-AUG-1987;

PR 25-APR-1988

PR 14-JUN-1990

PR 16-DEC-1991

PR 06-MAR-1992

PR 22-JUL-1992

PR 19-JAN-1994

PR 13-SEP-1994  
XX

PA (HEAL-) HEALTH RES INC.

PI Paolletti E;

DR WPI; 1999-493494/41.

PT Recombinant

PT other diseases such as influenza, Newcastle Disease and rabies.

PS Example 31; Fig 38; 163pp; English.

CC The present invention describes a recombinant poxvirus (I), comprising  
CC exogenous DNA encoding an antigenic determinant of a pathogen which is  
CC then expressed *in vivo* in infected host cells after administration to a  
CC patient and therefore induces an immunological response. (I) may be used  
CC to vaccinate patients against a wide range of diseases and disorders  
CC depending on the type of antigen encoded by the exogenous DNA. (I) may be  
CC used to vaccinate against diseases such as rabies, influenza and  
CC Newcastle disease. It is particularly useful for immunising against  
CC cancers. The poxvirus (I) also provides a means of manipulating  
CC lymphocytes and tumour cells for use in cell-based immunotherapeutic  
CC modalities for cancer. (I) also have enhanced safety compared to  
CC unattenuated viruses (attenuation reduces the virulence of the viruses)  
CC and known recombinant poxvirus vaccines. This increased level of safety  
CC reduces the possibility of a 'runaway' infection in the host and reduces  
CC the chance of transmission from vaccinated to unvaccinated individuals  
CC and contamination of the environment. The present sequence represents a  
CC murine p53 gene used in the exemplification of the present invention  
CC  
CC Sequence 1173 BP; 279 A; 349 G; 304 G; 241 T; 0 U; 0 Other;

Query Match	91.1%	Score 16.4	DB 2	Length 1173
Best Local Similarity	94.4%	Pred. No. 94		
Matches 17; Conservative	0	Mismatches	1	Indels 0
				Gaps 0

```

Oy      1 CGACTGTGATCCTCCAT 18
         |||||
Db      27 CGACTGTGACTCCTCCAT 10

```

## RESULT

ID ABS51008 standard; cDNA; 1208 BP.

AC ABS51008;

DT 21-OCT-2002 (first entry)

DE Mouse cDNA encoding bait protein tumour suppressor p53

KW Mouse; ss; gene; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;

KM protein-protein interaction map; PIM; anorectic; metabolic disorder.

**OS Mus musculus.**

PN WO200253726-

PD 11-JUL-2002.

PF 28-DEC-2001; 2001WO-EP015423

PR 02-JAN-2001; 2001US-0259377P

PA (HYBR-) HYBRIGENICS.

2000

XX  
X

DR P-PSDB; ABG69472.

PT Novel complex of protein-protein interactions in adipocyte cells for  
PT identifying compounds that modulate the protein-protein interactions and  
PT useful for treating obesity and metabolic disorders.

XX Claim 2, Page, 125pp; English.

XX The invention relates to a complex of protein-protein interactions  
CC (forming a protein-protein interaction map, PIM) in adipocyte cells as  
CC defined in the specification, or polynucleotides in adipocytes encoding  
CC for the polypeptides. Also included are a recombinant cell expressing the  
CC interacting polypeptides and a method of selecting a modulating compound  
CC in adipocyte cells, by cultivating a recombinant host cell on a selective  
CC medium containing a modulating compound and a reporter gene the  
CC expression of which is toxic for the recombinant host cell which is  
CC transformed with two vectors, where the first vector comprises a  
CC polynucleotide encoding a first hybrid polypeptide and DNA binding domain  
CC and the second vector comprising a polynucleotide encoding a second  
CC hybrid polypeptide and an activating domain that activates the toxic  
CC reporter gene, when the first and second hybrid polypeptides interact and  
CC selecting the modulating compound which inhibits the growth of the  
CC recombinant host cell (i.e. using the yeast two-hybrid system). The  
CC complexes are useful for identifying compounds that modulate the protein-  
CC protein interactions and useful for treating obesity and metabolic  
CC disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The  
CC compound isolated by the method is useful for treating and preventing  
CC obesity or metabolic diseases. The interactions between the proteins of  
CC the complex further define a set of selected interacting domains, SID.  
CC The present sequence encodes a member of the protein complex of the  
CC invention, used as the bait protein in the yeast two- hybrid assay. Note:  
CC The present sequence was not displayed in the specification but was  
CC obtained from its Genbank entry by the indexer

XX SQ Sequence 1208 BP; 283 A; 362 C; 314 G; 249 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 6; Length 1208;  
Best Local Similarity 94.4%; Pred. No. 94;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGGAATCCTCCAT 18  
DB 52 CGACTGGAATCCTCCAT 35

RESULT 12  
ADJ32551/C

ID ADJ32551 standard; DNA; 1208 BP.

AC ADJ32551;

DT 22-APR-2004 (first entry)

XX Mouse p53 DNA.

KW Integrin; cell proliferation; cancer; melanoma; adenoma; lymphoma;  
KW myeloma; carcinoma; glioma; plasmocytoma; sarcoma; thymoma; leukaemia;  
KW brain tumour; cytosarcoma; mouse; gene; ds.

XX Mus musculus.

XX Key Location/Qualifiers  
XX CDS 26..1198

XX /tag= a  
XX /product= "Mouse p53 protein"

XX US2003224993-A1.

XX 04-DEC-2003.

XX 17-MAR-2003; 2003US-00392113.

XX 12-OCT-2000; 2000US-0239705P.

XX 24-OCT-2000; 2000US-0242812P.  
XX 12-OCT-2001; 2001WO-US032127.

PR 15-MAR-2002; 2002US-0365078P.

XX (LAND/) LAND H.  
XX (DELE/) DELEU L.

XX Land H, Deleu L;

XX WPI; 2004-154528/15.

XX P-P8DB; ADJ32552.

PT Reducing the proliferation of a cancer cell involves inhibiting ligand  
PT binding to an integrin receptor on the cancer cell, where the integrin  
PT receptor comprises an integrin.

XX Disclosure; SEQ ID NO 32; 161pp; English.

XX The present invention relates to compositions and methods for reducing  
CC the proliferation of cancer cells through interaction with integrins. The  
CC invention is useful for reducing the proliferation of cancer cells e.g.  
CC melanoma, adenoma, lymphoma, myeloma, carcinoma, glioma, plasmocytoma,  
CC sarcoma, thymoma, leukaemia, skin cancer, retinal cancer, breast cancer,  
CC prostate cancer, colon cancer, esophageal cancer, stomach cancer,  
CC pancreas cancer, brain tumours, lung cancer, ovarian cancer, cervical  
CC cancer, hepatic cancer, gastrointestinal cancer, and head and neck cancer  
CC cells. The invention is also useful for identifying a therapeutic target  
CC which involves assaying potential reagent for activity. The present  
CC sequence is mouse p53 DNA.

XX SQ Sequence 1208 BP; 283 A; 362 C; 314 G; 249 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 12; Length 1208;  
Best Local Similarity 94.4%; Pred. No. 94;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGGAATCCTCCAT 18  
DB 52 CGACTGGAATCCTCCAT 35

RESULT 13  
AAQ85528/C

ID AAQ85528 standard; DNA; 396 BP.

AC AAQ85528;

DT 10-JAN-1996 (first entry)

XX DNA probe 34 detects DNA-protein complex in immortal cells.

KW DNA-protein complex; detection; proliferation; tumour formation;  
KW diagnose; malignancy; biopsy; probe; ss.

XX Synthetic.

XX WO9502701-A1.

XX 26-JAN-1995.

XX 13-JUL-1994; 94WO-EP002307.

XX 15-JUL-1993; 93DE-04323727.

XX (BOEF ) BOEHRINGER MANNHEIM GMBH.

XX Abken HJ, Albert W, Jungfer H;

XX WPI; 1995-067344/09.

PT New DNA-protein complex characteristic of cells with unlimited  
PT proliferation capacity - and its components and derived antibodies,  
PT useful in diagnosing malignant tumours.

XX Claim 1; Page 75; 106pp; German.

XX The DNA shown is found in human or animal cells that have an unlimited  
CC capacity for unlimited cell proliferation or tumor formation. They have  
CC no ability to promote immortalisation of the cells, and are usually found  
CC in a DNA-protein complex in the cell cytoplasm. The DNA is useful for  
CC detection of these complexes, and diagnosis of malignant tumours.  
CC Differentiation between cells with unlimited and only transient  
CC proliferation is possible. (See also AA085493-54)

XX  
SQ Sequence 396 BP; 105 A; 50 C; 138 G; 99 T; 0 U; 4 Other;  
Query Match 85.6%; Score 15.4; DB 2; Length 396;  
Best Local Similarity 94.1%; Pred. No. 3e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
|||  
Db 302 GACTGTGAATCCTCCAT 286

RESULT 14  
AAH05726  
ID AAH05726 standard; cDNA; 525 BP.  
AC AAH05726;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA clone (5'-primer) SEQ ID NO:2561.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUN-2000; 2000EP-00116126.  
XX  
PR 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
PT length cDNAs defined in the specification, and for the detection and/or  
PT diagnosis of the abnormality of the proteins encoded by the full-length  
PT cDNAs.  
XX  
PS Claim 1; SEQ ID NO 2561; 2537bp + Sequence Listing; English.  
XX  
XX The present invention describes primer sets for synthesizing 5602 full-  
CC length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification; where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention

XX  
SQ Sequence 525 BP; 84 A; 166 C; 151 G; 121 T; 0 U; 3 Other;  
Query Match 85.6%; Score 15.4; DB 4; Length 525;  
Best Local Similarity 94.1%; Pred. No. 3.1e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17  
|||  
Db 434 CGACTGTGAATCCTCCA 450

RESULT 15  
AAFI4040  
ID AAFI4040 standard; cDNA; 584 BP.  
XX  
AC AAFI4040;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Aspergillus oryzae EST SEQ ID NO:6563.  
XX  
KW Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX  
OS Aspergillus oryzae.  
XX  
PN WO200056762-A2.  
XX  
PD 28-SEP-2000.  
XX  
PF 22-MAR-2000; 2000WO-US007781.  
XX  
PR 22-MAR-1999; 99US-00273623.  
XX  
PA (NOVO) NOVO NORDISK BIOTECH INC.  
PA (NOVO) NOVO NORDISK AS.  
XX  
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
PI WPI; 2000-594572/56.  
XX  
DR  
XX  
PT Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags.  
XX  
PS Claim 88; Page 2686; 3161pp; English.  
XX  
XX The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring the  
CC global expression of genes from FF cells allows the production potential  
CC of the microorganisms to be improved. New genes may be discovered,  
CC possible functions of unknown open reading frames can be identified and  
CC gene copy number variation and stability can be monitored. The expression  
CC of genes can be used to study how FF cells adapt to changes in culture

CC conditions, environmental stress, spore morphogenesis, recombination,  
CC metabolic or catabolic pathway engineering. Using ESTs provides several  
CC advantages over genomic or random cDNA clones including elimination of  
CC redundancy as one spot on an array equals one gene or open reading frame,  
CC and organisation of the microarrays based on function of the gene  
CC products to facilitate analysis of the results. AAF07478 to AAF11247  
CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents  
CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from  
CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from  
CC *Trichoderma reesei*, which are all specifically claimed in the present  
CC invention  
XX  
SQ Sequence 584 BP; 148 A; 145 C; 137 G; 154 T; 0 U; 0 Other;  
Query Match 85.6%; Score 15.4; DB 3; Length 584;  
Best Local Similarity 94.1%; Pred. No. 3.1e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GACTGTGAATCCTCCAT 18  
DB 240 GATTGTGAATCCTCCAT 256  
RESULT 16  
AAF93528  
ID AAF93528 standard; cDNA; 608 BP.  
XX  
AC AAF93528;  
XX  
DT 21-MAY-2001 (first entry)  
XX  
DE cDNA encoding SRT protein isolated from prostate tissue SEQ ID 349.  
XX  
KM Human; SRT; gene therapy; gene mapping; tissue typing; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200107611-A2.  
XX  
PD 01-FEB-2001.  
XX  
PF 21-JUL-2000; 2000WO-US020006.  
XX  
PR 26-JUL-1999; 99US-0145701P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Goddard A, Wood WI;  
XX  
DR WPI; 2001-112729/12.  
XX  
PT New isolated nucleic acid molecule encoding a SRT polypeptide is useful  
PT for production of recombinant SRT polypeptides, gene mapping, diagnosing  
PT genetic disorders and for gene therapy.  
XX  
PS Claim 2; Fig 349; 663BP; English.  
XX  
CC Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding  
CC human SRT proteins. The cDNA sequences are isolated from various  
CC different human tissue cDNA libraries. The invention relates to a method  
CC for detecting cDNA encoding an SRT protein, a vector containing cDNA  
CC encoding SRT, a host cell transformed with the vector, an isolated SRT  
CC polypeptide, and an antibody which binds to SRT. The polynucleotide  
CC sequence can be used in gene therapy and is useful in the recombinant  
CC production of SRT polypeptides, as a hybridisation probe to screen  
CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to  
CC map the gene encoding the SRT polypeptides and analysing genetic  
CC disorders, tissue typing and disease tissue detection. The SRT  
CC polynucleotide sequences can be used in polymerase chain reaction,  
CC screening for new therapeutic molecules and generation of antisense RNA  
CC and DNA  
XX  
SQ Sequence 608 BP; 104 A; 181 C; 150 G; 167 T; 0 U; 6 Other;

Query Match 85.6%; Score 15.4; DB 5; Length 608;  
Best Local Similarity 94.1%; Pred. No. 3.1e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCCTCCA 17  
DB 299 CGACTGTGAATCCTCCA 315  
RESULT 17  
ADN62510/c  
ID ADN62510 standard; DNA; 795 BP.  
XX  
AC ADN62510;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE A. thaliana gene SEQ ID NO:15.  
XX  
KM ds; gene; seed size; organ size; plant; transgenic.  
XX  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..795  
FT /\*tag= a  
XX  
XX WO2003096797-A2.  
XX  
PD 27-NOV-2003.  
XX  
PE 14-MAY-2003; 2003WO-US014989.  
XX  
PR 15-MAY-2002; 2002US-0381100P.  
XX  
PA (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
PI Fernandes M, Xie Z, Dotson SB;  
XX  
DR WPI; 2004-042481/04.  
XX  
DR P-PSDB; ADN62511.  
XX  
PT Increasing seed and organ size of a plant by transforming the plant with  
PT a DNA construct comprising a promoter that functions in plants and  
PT selecting a desired plant from a population of transformed plants  
XX  
XX containing the DNA construct.  
XX  
PS Claim 1; SEQ ID NO 15; 71BP; English.  
XX  
CC The invention relates to a novel method for increasing seed and organ  
CC size of a plant by: transforming the plant with a DNA construct  
CC comprising a promoter that functions in plants, operably linked to a DNA  
CC molecule that encodes a protein; and selecting a desired plant from a  
CC population of transformed plants containing the DNA construct; where the  
CC desired plant exhibits increased seed and organ size compared to a plant  
CC of a same plant species not transformed to contain the DNA construct. The  
CC method of the invention is useful for increasing seed and organ size of a  
CC plant. The present sequence is used in the exemplification of the  
CC invention.  
XX  
SQ Sequence 795 BP; 220 A; 165 C; 199 G; 211 T; 0 U; 0 Other;  
Query Match 85.6%; Score 15.4; DB 12; Length 795;  
Best Local Similarity 94.1%; Pred. No. 3.2e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GACTGTGAATCCTCCAT 18  
DB 765 GACTGTGAATCCTCCAT 749  
RESULT 18



AAc44742  
ID AAc44742 standard; DNA; 806 BP.  
XX  
AC AAc44742;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Zea mays DNA fragment SEQ ID NO: 43964.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic; pathway;  
KW promoter; termination sequence; corn; ss.  
XX  
OS Zea mays subsp. mays.  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
PP 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0128645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0134449P.  
PR 30-APR-1999; 99US-0135048P.  
PR 30-APR-1999; 99US-0135407P.  
PR 04-MAY-1999; 99US-0135484P.  
PR 05-MAY-1999; 99US-0135485P.  
PR 06-MAY-1999; 99US-0135486P.  
PR 06-MAY-1999; 99US-0135487P.  
PR 07-MAY-1999; 99US-0135863P.  
PR 11-MAY-1999; 99US-0135865P.  
PR 14-MAY-1999; 99US-0135866P.  
PR 14-MAY-1999; 99US-0135867P.  
PR 14-MAY-1999; 99US-0135868P.  
PR 14-MAY-1999; 99US-0135869P.  
PR 18-MAY-1999; 99US-0135870P.  
PR 18-MAY-1999; 99US-0135871P.  
PR 19-MAY-1999; 99US-0135872P.  
PR 20-MAY-1999; 99US-0135873P.  
PR 21-MAY-1999; 99US-0135874P.  
PR 24-MAY-1999; 99US-0135875P.  
PR 25-MAY-1999; 99US-0135876P.  
PR 27-MAY-1999; 99US-0135877P.  
PR 28-MAY-1999; 99US-0135878P.  
PR 01-JUN-1999; 99US-0135879P.  
PR 03-JUN-1999; 99US-0135880P.  
PR 04-JUN-1999; 99US-0135881P.  
PR 07-JUN-1999; 99US-0135882P.  
PR 08-JUN-1999; 99US-0135883P.  
PR 10-JUN-1999; 99US-0135884P.  
PR 10-JUN-1999; 99US-0135885P.  
PR 14-JUN-1999; 99US-0135886P.  
PR 16-JUN-1999; 99US-0135887P.  
PR 17-JUN-1999; 99US-0135888P.  
PR 18-JUN-1999; 99US-0135889P.  
PR 18-JUN-1999; 99US-0135890P.  
PR 18-JUN-1999; 99US-0135891P.  
PR 18-JUN-1999; 99US-0135892P.  
PR 18-JUN-1999; 99US-0135893P.  
PR 18-JUN-1999; 99US-0135894P.  
PR 18-JUN-1999; 99US-0135895P.  
PR 18-JUN-1999; 99US-0135896P.  
PR 18-JUN-1999; 99US-0135897P.  
PR 18-JUN-1999; 99US-0135898P.  
PR 18-JUN-1999; 99US-0135899P.

PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 28-JUN-1999; 99US-0140695P.  
PR 29-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145066P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145092P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.

```

PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154032P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

```

```

Query March 85.6%; Score 15.4; DB 3; Length 806;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 GACTGTGAATCCTCCAT 18
Db 446 GACTGTGAATCCTCCAT 462

```

```

RESULT 19
AAK92361
ID AAK92361 standard; cDNA; 870 BP.
XX
AC AAK92361;
XX
DT 06-NOV-2001 (first entry)
XX

```

```

DE Human cDNA 5'-end sequence, SEQ ID NO: 821.
XX
XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
KW
XX Homo sapiens.
OS
XX EP1130094-A2.
PN
XX 05-SEP-2001.
PD
XX 07-JUL-2000; 2000EP-00114089.
PF
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
PS
XX Claim 2; SEQ ID NO 821; 1380pp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is the nucleotide sequence of the 5'-end of
CC a cDNA provided in the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in CD-
CC ROM format directly from EPO
XX
XX Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;
SQ

```

```

Query March 85.6%; Score 15.4; DB 4; Length 870;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CGACTGTGAATCCTCCA 17
Db 417 CGACTGTGAATCCTCCA 433

```

```

RESULT 20
AAK93899
ID AAK93899 standard; cDNA; 870 BP.
XX
AC AAK93899;
XX
DT 06-NOV-2001 (first entry)
XX

```

```

DE Human cDNA clone representative sequence, SEQ ID NO: 2359.
XX

```

```

KW Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX

```

```

OS Homo sapiens.
XX

```

```

XX EP1130094-A2.
XX

```

```

XX 05-SEP-2001.
XX

```

```

XX 07-JUL-2000; 2000EP-00114089.
XX

```

```

XX 08-JUL-1999; 99JP-00194486.
XX

```

PR 11-JAN-2000; 2000JP-00118774.  
 PR 02-MAY-2000; 2000JP-00183765.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX  
 DR WPI; 2001-524255/58.  
 XX  
 PT 830 Primers useful for synthesizing full length cDNA clones and their use  
 PT in genetic manipulation.  
 XX  
 PS Example 11; SEQ ID NO 2359; 1380bp + Sequence Listing; English.  
 CC  
 CC The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been isolated  
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have  
 CC been determined. Primers for synthesizing the full length cDNA are useful  
 CC for clarifying the function of the protein encoded by the cDNA. The full  
 CC length clones were obtained by construction of full length enriched cDNA  
 CC libraries that were synthesized by the oligo-capping method. The primers  
 CC enable the production of the full length cDNA easily without any special  
 CC methods. The present sequence was used as the representative sequence  
 CC from a human clone which was used in homology searches to identify the  
 CC clone. Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in CD-ROM format directly from  
 CC BPO  
 XX  
 SQ Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;  
 XX  
 Query Match 85.6%; Score 15.4; DB 4; Length 870;  
 Best Local Similarity 94.1%; Pred. No. 3.2e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 CGACTGTGATTCCTCCA 17  
 Db 417 CGACTGTGATTCCTCCA 433  
 XX  
 RESULT 21  
 ADL28788  
 ID ADL28788 standard; cDNA; 870 BP.  
 AC  
 AC ADL28788;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE 5' end of a human cDNA molecule SeqID 821.  
 XX  
 KW human; medicine; signal transduction; glycoprotein; transcription;  
 KW oligo-capping method; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1396543-A2.  
 XX  
 PD 10-MAR-2004.  
 XX  
 PP 07-JUL-2000; 2003EP-00025638.  
 XX  
 PR 08-JUL-1999; 99JP-00194486.  
 PR 11-JAN-2000; 2000JP-00118774.  
 PR 02-MAY-2000; 2000JP-00183865.  
 PR 07-JUL-2000; 2000EP-00114089.  
 XX  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX  
 DR WPI; 2004-204755/20.  
 XX

PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 PT length human cDNAs.  
 XX  
 XX  
 PS Disclosure; SEQ ID NO 821; 1340bp; English.  
 CC  
 CC This invention relates to a novel primers useful for synthesizing full  
 CC length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/  
 CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction.  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
 CC ends using the oligo-capping method. This polynucleotide sequence is the  
 CC 5' end of a full length human cDNA sequence of the invention.  
 XX  
 SQ Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;  
 XX  
 Query Match 85.6%; Score 15.4; DB 12; Length 870;  
 Best Local Similarity 94.1%; Pred. No. 3.2e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 CGACTGTGATTCCTCCA 17  
 Db 417 CGACTGTGATTCCTCCA 433  
 XX  
 RESULT 22  
 ADL30326  
 ID ADL30326 standard; cDNA; 870 BP.  
 AC  
 AC ADL30326;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE 3' end of a representative human cDNA cluster SeqID 2359.  
 XX  
 KW human; medicine; signal transduction; glycoprotein; transcription;  
 KW oligo-capping method; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1396543-A2.  
 XX  
 PD 10-MAR-2004.  
 XX  
 PP 07-JUL-2000; 2003EP-00025638.  
 XX  
 PR 08-JUL-1999; 99JP-00194486.  
 PR 11-JAN-2000; 2000JP-00118774.  
 PR 02-MAY-2000; 2000JP-00183865.  
 PR 07-JUL-2000; 2000EP-00114089.  
 XX  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX  
 DR WPI; 2004-204755/20.  
 XX  
 PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 PT length human cDNAs.  
 XX  
 XX  
 PS Example 18; SEQ ID NO 2359; 1340bp; English.  
 CC  
 CC This invention relates to a novel primers useful for synthesizing full  
 CC length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/  
 CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction.  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
 CC ends using the oligo-capping method. This polynucleotide sequence is the  
 CC 3' end of a representative human DNA cluster of the invention.

XX SQ Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;  
Query Match 85.6%; Score 15.4; DB 12; Length 870;  
Best Local Similarity 94.1%; Pred. No. 3.2e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 CGACTGTGATCTCTCCA 17  
|||||||  
DB 417 CGACTGTGATCTCTCCA 433  
RESULT 23  
AAC40009/c  
ID AAC40009 standard; DNA; 1077 BP.  
XX AC AAC40009;  
XX 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26716.  
XX KW Hybridisation assay; genetic mapping; gene expression control;  
XX KW protein identification; signal transduction pathway; metabolic pathway;  
XX KW promoter; termination sequence; ss.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-012180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135625P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.

```

PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149358P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155658P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158028P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159328P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

```

```

Query Match      85.6%; Score 15.4; DB 3; Length 1077;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy 2 GACGTGATCTCTCCAT 18
Db 825 GACGTGATCTCTCCAT 809

RESULT 24
AAH14165
ID AAH14165 standard; cDNA, 1364 BP.
XX
AC AAH14165;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SFG ID NO:11397.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUN-2000; 2000EP-00116126.
XX
PR 29-JUN-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 11397; 2537bp + Sequence listing; English.
XX
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH1632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1364 BP; 267 A; 415 C; 370 G; 312 T; 0 U; 0 Other;

```

```

Query Match      85.6%; Score 15.4; DB 4; Length 1364;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CGACTGTGAATCCTCCA 17  
 |||||  
 DB 434 CGACTGTGAATCCTCCA 450

## RESULT 25

ADD69659  
 ID ADD69659 standard; cDNA, 1487 BP.

AC ADD69659;

DT 15-JAN-2004 (first entry)

DE Human REMAP cDNA - SEQ ID 88.

XX human; receptor and membrane-associated protein; REMAP; cytosolic;  
 KW antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective;  
 KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory;  
 KW thymometric; cell proliferative; cancer; atherosclerosis; neurological;  
 KW epilepsy; Huntington's disease; stroke; immune; inflammatory; AIDS;  
 KW allergy; developmental; hypothyroidism; Cushing's syndrome; infection;  
 KW ss; gene.

XX Homo sapiens.

PN WO2003048305-A2.

PD 12-JUN-2003.

PF 13-NOV-2002; 2002WO-US036759.

PR 13-NOV-2001; 2001US-0333097P.

PR 15-NOV-2001; 2001US-0335274E.

PR 14-DEC-2001; 2001US-0340542P.

PR 18-DEC-2001; 2001US-0342166P.

PR 11-JAN-2002; 2002US-0347580P.

PR 14-JAN-2002; 2002US-0348687P.

XX (INCYTE-) INCYTE GENOMICS INC.

PI Duggan BM, Yang J, Gietzen KJ, Lee SY, Tang YT, Azimzal Y;

PI Chawla NK, Warren BA, Barroso I, Becha SD, Yue H, Lehr-Mason PM;

PI Thangavelu K, Lee S, Emerling BM, Kable AE, Khare R, Baughn MR;

PI Gandhi AR, Tran UK, Richardson TW, Margis JP, Lal PG, Forsythe LJ;

PI Lee EA, Swarnakar A, Kallick DA, Griffin JA, Elliott VS, Gorvay AE;

PI Hatalia AA, Ison CH, Jin P, Jiang X, Jackson AA, Bhatia U;

PI Buttrill JD, Blake UT, Ho A, Zheng W, Gao J;

XX WPI; 2003-513744/48.

DR P-PSDB; ADD69612.

XX New human receptors and membrane-associated proteins (REMAP), useful for

PT diagnosing, treating and preventing diseases or conditions associated

PT with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or

PT infections.

XX Claim 5; SEQ ID NO 88; 298bp; English.

XX The invention relates to a novel isolated polypeptide comprising a human

XX receptor and membrane-associated protein (REMAP). The polypeptide of the

XX invention demonstrates cytosolic, antiarteriosclerotic, anticonvulsant,

XX neurotropic, neuroprotective, cerebroprotective, anti-HIV, antiinflammatory,

XX antiinflammatory and thymometric activities and may be useful for

XX treating and diagnosing various disorders including those which are cell

XX proliferative such as cancer and atherosclerosis, neurological including

XX epilepsy, Huntington's disease and stroke, immune/inflammatory

XX particularly AIDS and allergies and developmental such as hypothyroidism

XX and Cushing's syndrome, as well as infections. The current sequence is

XX that of the human REMAP cDNA of the invention.

XX Sequence 1487 BP; 319 A; 442 C; 392 G; 334 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 1487;  
 Best Local Similarity 94.1%; Pred. No. 3.3e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17  
 |||||  
 DB 542 CGACTGTGAATCCTCCA 558

## RESULT 26

AAK94874  
 ID AAK94874 standard; cDNA, 1490 BP.

AC AAK94874;

DT 06-NOV-2001 (first entry)

DE Human full-length cDNA, SEQ ID NO: 4064.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

PN EP130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-00114089.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

PI Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

DR P-PSDB; AAK93912.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

PT in genetic manipulation.

XX Claim 8; SEQ ID NO 4064; 1380bp + Sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA

XX clones. 830 cDNA molecules encoding a human protein have been isolated

XX and nucleotide sequences of 5' and 3' ends of the cDNA molecules have

XX been determined. Primers for synthesizing the full length cDNA are useful

XX for clarifying the function of the protein encoded by the cDNA. The full

XX length clones were obtained by construction of full length enriched cDNA

XX libraries that were synthesized by the oligo-capping method. The primers

XX enable the production of the full length cDNA easily without any special

XX methods. The present sequence is a full length human cDNA of the

XX invention. Note: The sequence data for this patent did not form part of

XX the printed specification, but was obtained in CD-ROM format directly

XX from EPO

XX Sequence 1490 BP; 288 A; 455 C; 399 G; 348 T; 0 U; 0 Other;

XX Query Match 85.6%; Score 15.4; DB 4; Length 1490;

XX Best Local Similarity 94.1%; Pred. No. 3.3e+02;

XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17

DB 417 CGACTGTGAATCCTCCA 433

RESULT 27

ADL32031

ID ADL32031 standard; cDNA, 1490 BP.

XX ADL32031;  
AC 20-MAY-2004 (first entry)  
XX  
XX Full length human cDNA clone Segid 4064.  
DE  
XX human; medicine; signal transduction; glycoprotein; transcription;  
KM oligo-capping method; ss; gene.  
XX  
XX Homo sapiens.  
OS  
XX BP1396543-A2.  
PN  
XX 10-MAR-2004.  
PD  
XX 07-JUL-2000; 2003EP-00025638.  
PF  
XX 08-JUL-1999; 99JP-00194486.  
PR 11-JAN-2000; 2000JP-00118774.  
PR 02-MAY-2000; 2000JP-00183865.  
PR 07-JUL-2000; 2000EP-00114089.  
XX  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ichii S, Kawai Y,  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX WPI; 2004-204755/20.  
DR P-PSDB; ADL32032.  
XX  
XX New oligonucleotide primers (830 CDNAs) useful for synthesizing full  
PT length human cDNAs.  
XX  
XX Example 1; SEQ ID NO 4064; 1340bp; English.  
XX  
XX This invention relates to a novel primers useful for synthesizing full  
CC length cDNA molecules that encode human proteins. Specifically, it refers  
CC to secretory or membrane proteins that are potential therapeutic agents/  
CC target molecules in the field of medicine, and in particular genes  
CC encoding proteins that are associated with signal transduction,  
CC glycoproteins and transcription. The present invention describes a method  
CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
CC ends using the oligo-capping method. This polynucleotide sequence is a  
CC full length human cDNA clone of the invention.  
XX  
XX Sequence 1490 BP; 288 A; 455 C; 399 G; 348 T; 0 U; 0 Other;  
SQ  
XX  
XX Query Match 85.6%; Score 15.4; DB 12; Length 1490;  
Best Local Similarity 94.1%; Pred. No. 3.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCTCTCA 17  
Db 417 CGACTGTGATTCCTCCA 433  
XX  
XX RESULT 28  
XX AAH64732  
ID AAH64732 standard; cDNA; 1510 BP.  
XX  
XX AC AAH64732;  
XX  
XX 11-SEP-2001 (first entry)  
XX  
XX Human secreted protein cDNA, SEQ ID NO: 8.  
DE  
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
KM GENSET; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200142451-A2.  
PN

XX 14-JUN-2001.  
PD  
XX 07-DEC-2000; 2000WO-IB001938.  
PF  
XX 08-DEC-1999; 99US-0169629P.  
PR 06-MAR-2000; 2000US-0187470P.  
XX  
XX (GENSET) GENSET.  
XX  
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;  
PI WPI; 2001-367870/38.  
DR P-PSDB; AAG89129.  
XX  
XX Full length GENSET human nucleic acids encoding potentially secreted  
PT proteins, useful in gene therapy and vaccination against a variety of  
PT diseases, and for diagnosis of those diseases.  
XX  
XX Claim 7; Page 566-568; 921pp; English.  
XX  
XX The invention relates to full length GENSET human nucleic acids encoding  
CC potentially secreted proteins. The nucleic acids and the polypeptides  
CC they encode may be used in the prevention, treatment and diagnosis of  
CC diseases associated with inappropriate GENSET gene expression. For  
CC example, they be used to treat disorders associated with decreased GENSET  
CC gene expression by rectifying mutations or deletions in a patient's  
CC genome that affect the activity of GENSET or by supplementing the  
CC patient's own production of GENSET polypeptides. Conversely, antisense  
CC nucleic acid molecules may be administered to down regulate GENSET  
CC expression by binding with the cells' own genes and preventing their  
CC expression. The sense and antisense nucleic acids may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar nucleic acid sequences in samples, and hence to determine which  
CC patients may be in need of restorative therapy. The GENSET polypeptides  
CC may be used as antigens in the production of antibodies and in assays to  
CC identify modulators (agonists and antagonists) of GENSET polypeptide  
CC expression and activity. The present sequence is a GENSET nucleic acid of  
CC the invention  
XX  
XX Sequence 1510 BP; 307 A; 452 C; 404 G; 347 T; 0 U; 0 Other;  
SQ  
XX  
XX Query Match 85.6%; Score 15.4; DB 5; Length 1510;  
Best Local Similarity 94.1%; Pred. No. 3.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGACTGTGAATCTCTCA 17  
Db 405 CGACTGTGATTCCTCCA 421  
XX  
XX RESULT 29  
XX AAH64771  
ID AAH64771 standard; cDNA; 1523 BP.  
XX  
XX AC AAH64771;  
XX  
XX 11-SEP-2001 (first entry)  
XX  
XX Human secreted protein cDNA, SEQ ID NO: 47.  
DE  
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
KM GENSET; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200142451-A2.  
PN 14-JUN-2001.  
PD  
XX 07-DEC-2000; 2000WO-IB001938.  
PF  
XX 08-DEC-1999; 99US-0169629P.  
PR

```

PR 06-MAR-2000; 2000US-0187470P.
XX
XX (GENSET ) GENSET.
XX
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
XX WPI; 2001-367870/38.
XX
XX P-PSDB; AAG89168.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
XX proteins, useful in gene therapy and vaccination against a variety of
XX diseases, and for diagnosis of those diseases.
XX
XX Claim 7; Page 612-614; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased GENSET
XX gene expression by rectifying mutations or deletions in a patient's
XX genome that affect the activity of GENSET or by supplementing the
XX patient's own production of GENSET polypeptides. Conversely, antisense
XX nucleic acid molecules may be administered to down regulate GENSET
XX expression by binding with the cells' own genes and preventing their
XX expression. The sense and antisense nucleic acids may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acid sequences in samples, and hence to determine which
XX patients may be in need of restorative therapy. The GENSET polypeptides
XX may be used as antigens in the production of antibodies and in assays to
XX identify modulators (agonists and antagonists) of GENSET polypeptide
XX expression and activity. The present sequence is a GENSET nucleic acid of
XX the invention
XX
SQ Sequence 1523 BP; 308 A; 462 C; 404 G; 349 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1523;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCCTCCA 17
Db 432 CGACTGTGATCCTCCA 448
RESULT 30
AAV43615
ID AAV43615 standard; DNA; 1535 BP.
XX
XX AAV43615;
XX
XX 24-SEP-1998 (first entry)
XX
XX Human secreted protein 15 encoding DNA.
XX
XX Secreted protein; human; cell proliferation; cytokine activity;
XX tissue growth; cellular differentiation; regeneration; activin; inhibin;
XX chemotactic; haemostatic; thrombolytic; tumour inhibition;
XX anti-inflammatory activity; biomarker; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 220..1224
XX /tag= a
XX /product= "human secreted protein"
XX
XX WO9825959-A2.
XX
XX 18-JUN-1998.
XX
XX 11-DEC-1997; 97WO-US022787.
XX

```

```

PR 11-DEC-1996; 96US-0032757P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Escobedo J, Hu Q, Garcia P, Williams LT, Kothakota S;
XX
XX WPI; 1998-348453/30.
XX
XX P-PSDB; AAM63695.
XX
XX Secreted human polypeptides - having cytokine, cell proliferation or
XX differentiation, activin or inhibin, tumour inhibition or anti-
XX inflammatory activities.
XX
XX Claim 6; Page 42; 78pp; English.
XX
XX This DNA encodes a human secreted protein. The specification provides
XX secreted protein sequences (AAM63681 to AAM63699) encoded by the nucleic
XX acid sequences shown in AAV43601 to AAV43619. The invention provides a
XX method of identifying a secreted polypeptide which is modified by rough
XX microsome. The secreted proteins can be used in assays to determine
XX biological activities, such as cytokine, cell proliferation, or cellular
XX differentiation activities, tissue growth or regeneration, activin or
XX inhibin activity, chemotactic or chemokinetic activity, haemostatic or
XX thrombolytic activity, receptor/ligand activity, tumour inhibition, or
XX anti-inflammatory activity. The proteins can also be used as biomarkers,
XX to identify tissues or cell types which express the proteins, or a stage-
XX or disease-specific alteration in protein expression. They can be used in
XX protein interaction assays, to identify ligands or binding proteins.
XX Compounds which affect the biological activities of the secreted proteins
XX or their ability to interact with specific ligands can be identified
XX using the proteins in screening assays. The proteins and antibodies that
XX bind specifically to the protein can also be used to design diagnostic
XX tests and therapeutic compositions for diseases which may be associated
XX with altered expression of these proteins. Fusion proteins comprising,
XX e.g. signal sequences or transmembrane domains of the proteins can be
XX used to target other protein domains to cellular membrane or they can be
XX secreted extracellularly
XX
SQ Sequence 1535 BP; 313 A; 463 C; 408 G; 351 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 2; Length 1535;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCCTCCA 17
Db 435 CGACTGTGATCCTCCA 451
RESULT 31
ABZ68515
ID ABZ68515 standard; DNA; 1664 BP.
XX
XX ABZ68515;
XX
XX 22-APR-2003 (first entry)
XX
XX Nucleotide sequence of human CGLI polypeptide.
XX
XX Human; congenital generalized lipodystrophy protein; CGLI; 11q13;
XX chromosome 11; congenital generalized lipodystrophy; lipodystrophy;
XX diabetes; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 345..1541
XX /tag= a
XX /product= "GLI"
XX
XX FR82824332-A1.
XX
XX 08-NOV-2002.
XX

```



```

XX 04-MAY-2001; 2001FR-00006037.
XX
XX 04-MAY-2001; 2001FR-00006037.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX (NAGE-) CENT NAT GENOTYPAGE.
XX
XX Magre J, Capeau J, Lathrop M, Delphine M;
XX
XX WPI; 2003-142459/14.
XX P-PSDB; ABP97511.
XX
XX Nucleic acid encoding a congenital generalized lipodystrophy gene cgll
XX PT and mutations of that gene, useful to prevent and treat congenital
XX PT generalized lipodystrophy and obesity.
XX
XX Claim 4; Page 11; 115pp; French.
XX
XX The present sequence encodes a human congenital generalized lipodystrophy
XX CC protein, designated CGLI. The CGLI gene is localised at 11q13 on
XX CC chromosome 11. CGLI is responsible for congenital generalized
XX CC lipodystrophy. CGLI polypeptides and polynucleotides are used for
XX CC preventing or treating lipodystrophy or diabetes. CGLI polypeptides are
XX CC also useful as immunogens for raising antibodies.
XX
XX Sequence 1664 BP; 382 A; 481 C; 428 G; 373 T; 0 U; 0 Other;
XX
XX Query Match 85.6%; Score 15.4; DB 8; Length 1664;
XX Best Local Similarity 94.1%; Pred. No. 3.4e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 CGACTGTGATCTCTCCA 17
XX Db 560 CGACTGTGATCTCTCCA 576
XX
XX RESULT 32
XX ID ADP18659 standard; cDNA; 1664 BP.
XX
XX ADP18659;
XX
XX 26-AUG-2004 (first entry)
XX
XX Human TAT417 cDNA used to treat cancer Segid 19.
XX
XX tumour; growth inhibitory; cytotoxic; cytostatic; gene therapy; cancer;
XX KW human; TAT; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2004045516-A2.
XX
XX 03-JUN-2004.
XX
XX 13-NOV-2003; 2003WO-US036298.
XX
XX 15-NOV-2002; 2002US-0426847P.
XX PR 06-DEC-2002; 2002US-0431250P.
XX PR 31-DEC-2002; 2002US-0437344P.
XX
XX (GERTH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Frantz G, Goddard A, Gonzalez L, Gurney AL;
XX PI Polakis P, Polson A, Wood WI, Wu TD, Zhang Z;
XX
XX WPI; 2004-420515/39.
XX P-PSDB; ADP18680.
XX
XX New antibody that binds to tumor-associated antigenic target polypeptide
XX PT (TAT), useful in preparing a composition for diagnosing or treating
XX PT tumor.

```

```

XX Claim 1; SEQ ID NO 19; 183pp; English.
XX
XX This invention relates to novel isolated antibodies that are useful for
XX CC the diagnosis and treatment of a tumour. Specifically, it refers to DNA
XX CC that is detectably labelled and conjugated to a growth inhibitory or
XX CC cytotoxic agent and comprises toxins, antibiotics, radioactive isotopes
XX CC or nucleolytic enzymes. The present invention refers to the toxin
XX CC maytansinoid or calicheamicin, where the host cell is one of chinese
XX CC hamster ovary (CHO), yeast or Escherichia coli cell. Furthermore, these
XX CC pharmaceutical compositions are cytostatic and can be used for gene
XX CC therapy purposes to treat various cancers. This polynucleotide sequence
XX CC is a human TAT cDNA sequence of the invention.
XX
XX Sequence 1664 BP; 382 A; 481 C; 428 G; 373 T; 0 U; 0 Other;
XX
XX Query Match 85.6%; Score 15.4; DB 12; Length 1664;
XX Best Local Similarity 94.1%; Pred. No. 3.4e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 CGACTGTGATCTCTCCA 17
XX Db 560 CGACTGTGATCTCTCCA 576
XX
XX RESULT 33
XX ID AAV55741 standard; cDNA; 1755 BP.
XX
XX AAV55741;
XX
XX 23-MAR-1999 (first entry)
XX
XX Human secreted protein clone as294_3 coding sequence.
XX
XX Secreted protein; human; nutritional supplements; immune stimulant;
XX KW immune suppressor; haematopoiesis regulator; tissue growth; chemotaxis;
XX KW activin/inhibin; chemokinesis; haemostasis; thrombolysis;
XX KW receptor/ligand activity; anti-inflammatory; tumour inhibitor;
XX KW cadherin/tumour invasion suppressor; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT CDS 266..1654
XX FT /*tag= a
XX
XX MO9855614-A2.
XX
XX 10-DEC-1998.
XX
XX 01-JUN-1998; 98WO-US011210.
XX
XX 04-JUN-1997; 97US-00868696.
XX PR 04-JUN-1997; 97US-00868697.
XX PR 04-JUN-1997; 97US-00868698.
XX PR 04-JUN-1997; 97US-00868699.
XX PR 04-JUN-1997; 97US-00868900.
XX PR 04-JUN-1997; 97US-00869191.
XX PR 04-JUN-1997; 97US-00869192.
XX PR 04-JUN-1997; 97US-00869193.
XX PR 04-JUN-1997; 97US-00869194.
XX PR 29-MAY-1998; 98US-0087252P.
XX
XX (GENV ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Treacy M, Spaulding V;
XX PI Agostino MJ, Howes SR, Fenchel K;
XX
XX WPI; 1999-059912/05.
XX P-PSDB; AAW73623.
XX

```

PT New polynucleotides encoding secreted human proteins - derived from human  
 PT foetal brain, adult brain, foetal kidney, adult ovary, adult retina,  
 PT adult placenta or adult uterus cDNA libraries.

XX Claim 1; Page 71-72; 127pp; English.

CC This sequence encodes a human secreted protein of the invention. This DNA  
 CC sequence was isolated from a human foetal brain cDNA library, and was  
 CC designated clone as294\_3. The DNAs and proteins are predicted to have  
 CC biological activities which would make them suitable for treating,  
 CC preventing or ameliorating medical conditions in humans and animals,  
 CC although no supporting data is given. Suggested activities include  
 CC nutritional sources or supplements, immune stimulating or suppressing  
 CC activity, haematopoiesis regulating activity, tissue growth activity,  
 CC activ/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, and tumour  
 CC inhibition activity. The DNAs are also stated to be useful for gene  
 CC therapy. A host cell transfected with the DNA, or its subfragments and  
 CC variants is useful for recombinant production of the human secreted  
 CC protein clones

XX Sequence 1755 BP; 384 A; 519 C; 452 G; 399 T; 0 U; 1 Other;

Query Match 85.6%; Score 15.4; DB 2; Length 1755;  
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGATTCCTCCA 17  
 |||||  
 DB 673 CGACTGTGATTCCTCCA 689

RESULT 34

ABO92067 ABO92067 standard; cDNA, 1755 BP.

XX ABO92067;

DT 04-OCT-2002 (first entry)

XX Human polynucleotide SEQ ID NO 64.

Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;  
 antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;  
 antidiabetic; antidiabetic; antidiabetic; antidiabetic;  
 immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 stem cell; growth factor; nervous system disease; neuropathy;  
 Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 osteoporosis; severe combined immunodeficiency; SCID; infection;  
 multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.

XX Homo sapiens.

XX US2002065394-A1.

XX 30-MAY-2002.

XX 22-DEC-2000; 2000US-00745763.

XX 18-MAR-1998; 98US-00040963.

XX (JACO/) JACOBS K.  
 PA (MCCO/) MCCOY J M.  
 PA (LAVA/) LAVALLIE E R.  
 PA (COLL/) COLLINS-RACIE L A.  
 PA (EVAN/) EVANS C.  
 PA (MERB/) MERBERG D.  
 PA (TREA/) TREACY M.  
 PA (SPAU/) SPAULDING V.

PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Spaulding V;  
 XX WPI: 2002-582343/62.  
 DR P-PSDB; ABP61854.

PT Novel secreted or transmembrane protein and polynucleotide encoding the  
 PT protein, useful for diagnosis and treatment of neurological disorders,  
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.  
 XX Claim 227; Page 219-220; 284pp; English.

CC The invention relates to human secreted or transmembrane protein (I),  
 CC their fragments and is encoded by specific complementary deoxyribonucleic  
 CC acid (cDNA) inserts (II), where the protein is substantially free from  
 CC other mammalian proteins. (II) are useful for preventing, treating or  
 CC ameliorating a medical condition, especially immunological treatment or  
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,  
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,  
 CC stem cell growth factor activity and activin or inhibin-related  
 CC activities. (I) can be used to manipulate stem cells in culture to give  
 CC rise to neuroepithelial cells that can be used to augment or replace  
 CC cells damaged by illness, autoimmune disease, accidental damage or  
 CC genetic disorders. (I) induces the proliferation of neural cells and  
 CC regeneration of nerve and brain tissue and is useful for the treatment of  
 CC central and peripheral nervous system diseases and neuropathies, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis. (II) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,  
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 CC periodontal disease. (I) is also useful for gut protection or  
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (II) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polynucleotide of the invention

XX Sequence 1755 BP; 384 A; 519 C; 452 G; 399 T; 0 U; 1 Other;

Query Match 85.6%; Score 15.4; DB 6; Length 1755;  
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGATTCCTCCA 17  
 |||||  
 DB 673 CGACTGTGATTCCTCCA 689

RESULT 35

AAF21834 AAF21834 standard; DNA, 1892 BP.

XX AAF21834;

DT 27-MAR-2001 (first entry)

XX Human breast and ovarian cancer associated antigen gene SEQ ID 221.

XX Human; breast cancer; ovarian cancer; cytosolic; immunosuppressive;  
 XX neurotropic; neuroprotective; antiviral; antiinflammatory; hepatotropic;  
 XX antidiabetic; antiinflammatory; antidiabetic; antidiabetic;  
 XX antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
 XX Addison's disease; allergy; autoimmune haemolytic anaemia;  
 XX autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
 XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
 XX cardiovascular disorder; wound healing; neurological disease; da.

XX Homo sapiens.  
 OS  
 XX  
 PN WO200055173-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PP 08-MAR-2000; 2000WO-US005881.  
 XX  
 PR 12-MAR-1999; 99US-0124270P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI; 2000-611515/58.  
 XX  
 DR P-PSDB; AAB58931.  
 XX  
 PT New human breast and ovarian cancer associated gene sequences and the  
 PT polypeptides encoded by these genes, useful in the prevention, treatment  
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
 PT neurological diseases.  
 XX  
 PS Claim 1; Page 647-648; 1299pp; English.  
 XX  
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
 CC associated with breast and ovarian cancer. Included in the invention are  
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterisation of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; nocrotic;  
 CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;  
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and  
 CC protein sequences are used in the diagnosis of cancer, particularly  
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
 CC and agonists may also be used in the diagnosis, prevention and treatment  
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC cardiovascular disorders such as myocardial ischaemia; wound healing;  
 CC neurological diseases such as cerebral anoxia and epilepsy; and  
 CC infectious diseases  
 XX  
 SQ Sequence 1892 BP; 441 A; 553 C; 478 G; 417 T; 0 U; 3 Other;  
 Query Match 85.6%; Score 15.4; DB 3; Length 1892;  
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTGAATCTCTCA 17  
 DB 763 CGACTGTGAATCTCTCA 779  
 RESULT 36  
 ABL19347/c  
 ID ABL19347 standard; DNA; 2148 BP.  
 XX  
 AC ABL19347;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9514.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.

PD 27-SEP-2001.  
 XX  
 PP 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 XX  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PERE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Claim 1; SEQ ID NO 9514; 21pp + Sequence listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutic and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (AAB57737-  
 CC AAB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 2148 BP; 534 A; 559 C; 506 G; 549 T; 0 U; 0 Other;  
 Query Match 85.6%; Score 15.4; DB 4; Length 2148;  
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTGAATCTCTCA 17  
 DB 444 CGACTGTGAATCTCTCA 428  
 RESULT 37  
 ABL15659/c  
 ID ABL15659 standard; CDNA; 2243 BP.  
 XX  
 AC ABL15659;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41459.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 EN  
 XX  
 PD 27-SEP-2001.  
 XX  
 PP 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 XX  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PERE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 XX  
 DR P-PSDB; AAB71556.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.

XX Claim 1; SEQ ID NO 41459; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 2243 BP; 541 A; 594 C; 534 G; 574 T; 0 U; 0 Other;

XX Query Match 85.6%; Score 15.4; DB 4; Length 2243;

XX Best Local Similarity 94.1%; Pred. No. 3.5e+02;

XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCA 17  
Db 539 CGACAGTGAATCCTCCA 523

RESULT 38

AB102040/C  
ID ABL02040 standard; cDNA; 3703 BP.

XX ABL02040;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 602.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE ) PE CORP NY.

PI Venter UC, Adams M, Li FWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB57937.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.

XX Claim 1; SEQ ID NO 602; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 3703 BP; 1041 A; 814 C; 839 G; 1009 T; 0 U; 0 Other;

XX Query Match 85.6%; Score 15.4; DB 4; Length 3703;

XX Best Local Similarity 94.1%; Pred. No. 3.6e+02;

XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCA 17  
Db 3502 CGACAGTGAATCCTCCA 3486

RESULT 39

AAK75168  
ID AAK75168 standard; DNA; 4357 BP.

XX AAK75168;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29980.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytosolic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US001354.

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0198874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225271P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225278P.

PR 14-AUG-2000; 2000US-0225279P.

PR 14-AUG-2000; 2000US-0225281P.

PR 14-AUG-2000; 2000US-0225282P.

PR 14-AUG-2000; 2000US-0225283P.

PR 14-AUG-2000; 2000US-0225284P.

PR 14-AUG-2000; 2000US-0225285P.

PR 14-AUG-2000; 2000US-0225286P.

PR 14-AUG-2000; 2000US-0225287P.

PR 14-AUG-2000; 2000US-0225288P.

PR 14-AUG-2000; 2000US-0225289P.

PR 14-AUG-2000; 2000US-0225290P.

PR 14-AUG-2000; 2000US-0225291P.

PR 14-AUG-2000; 2000US-0225292P.

PR 14-AUG-2000; 2000US-0225293P.

PR 14-AUG-2000; 2000US-0225294P.

PR 14-AUG-2000; 2000US-0225295P.

PR 14-AUG-2000; 2000US-0225296P.

PR 14-AUG-2000; 2000US-0225297P.

PR 14-AUG-2000; 2000US-0225298P.

PR 14-AUG-2000; 2000US-0225299P.

PR 14-AUG-2000; 2000US-0225300P.

PR 14-AUG-2000; 2000US-0225301P.

PR 14-AUG-2000; 2000US-0225302P.

PR 14-AUG-2000; 2000US-0225303P.

PR 14-AUG-2000; 2000US-0225304P.

PR 14-AUG-2000; 2000US-0225305P.

PR 14-AUG-2000; 2000US-0225306P.

PR 14-AUG-2000; 2000US-0225307P.

PR 14-AUG-2000; 2000US-0225308P.

PR 14-AUG-2000; 2000US-0225309P.

PR 14-AUG-2000; 2000US-0225310P.

PR 14-AUG-2000; 2000US-0225311P.

PR 14-AUG-2000; 2000US-0225312P.

PR 14-AUG-2000; 2000US-0225313P.

PR 14-AUG-2000; 2000US-0225314P.

PR 14-AUG-2000; 2000US-0225315P.

PR 14-AUG-2000; 2000US-0225316P.

PR 14-AUG-2000; 2000US-0225317P.

PR 14-AUG-2000; 2000US-0225318P.

PR 14-AUG-2000; 2000US-0225319P.

PR 14-AUG-2000; 2000US-0225320P.

PR 14-AUG-2000; 2000US-0225321P.

PR 14-AUG-2000; 2000US-0225322P.

PR 14-AUG-2000; 2000US-0225323P.

PR 14-AUG-2000; 2000US-0225324P.

PR 14-AUG-2000; 2000US-0225325P.

PR 14-AUG-2000; 2000US-0225326P.

PR 14-AUG-2000; 2000US-0225327P.

PR 14-AUG-2000; 2000US-0225328P.

PR 14-AUG-2000; 2000US-0225329P.

PR 14-AUG-2000; 2000US-0225330P.

PR 14-AUG-2000; 2000US-0225331P.

PR 14-AUG-2000; 2000US-0225332P.

PR 14-AUG-2000; 2000US-0225333P.

PR 14-AUG-2000; 2000US-0225334P.

PR 14-AUG-2000; 2000US-0225335P.

PR 14-AUG-2000; 2000US-0225336P.

PR 14-AUG-2000; 2000US-0225337P.

PR 14-AUG-2000; 2000US-0225338P.

PR 14-AUG-2000; 2000US-0225339P.

PR 14-AUG-2000; 2000US-0225340P.

PR 14-AUG-2000; 2000US-0225341P.

PR 14-AUG-2000; 2000US-0225342P.

PR 14-AUG-2000; 2000US-0225343P.

PR 14-AUG-2000; 2000US-0225344P.

PR	17-NOV-2000	2000US-02295455
PR	05-SEP-2000	2000US-02295097
PR	05-SEP-2000	2000US-02295139
PR	06-SEP-2000	2000US-02304376
PR	06-SEP-2000	2000US-02304387
PR	08-SEP-2000	2000US-02312429
PR	08-SEP-2000	2000US-02312439
PR	08-SEP-2000	2000US-02312444
PR	08-SEP-2000	2000US-02314133
PR	08-SEP-2000	2000US-02314134
PR	08-SEP-2000	2000US-02320807
PR	08-SEP-2000	2000US-02320818
PR	12-SEP-2000	2000US-02331968
PR	14-SEP-2000	2000US-02332976
PR	14-SEP-2000	2000US-02332988
PR	14-SEP-2000	2000US-02332997
PR	14-SEP-2000	2000US-02334233
PR	14-SEP-2000	2000US-02344747
PR	14-SEP-2000	2000US-02344749
PR	14-SEP-2000	2000US-02344774
PR	25-SEP-2000	2000US-02349979
PR	25-SEP-2000	2000US-02349987
PR	26-SEP-2000	2000US-02353844
PR	26-SEP-2000	2000US-02353846
PR	27-SEP-2000	2000US-02355343
PR	27-SEP-2000	2000US-02355367
PR	29-SEP-2000	2000US-02363277
PR	29-SEP-2000	2000US-02363279
PR	29-SEP-2000	2000US-02363686
PR	29-SEP-2000	2000US-02363688
PR	29-SEP-2000	2000US-02365399
PR	29-SEP-2000	2000US-02365702
PR	02-OCT-2000	2000US-02370377
PR	02-OCT-2000	2000US-02370387
PR	02-OCT-2000	2000US-02370396
PR	02-OCT-2000	2000US-02370397
PR	02-OCT-2000	2000US-02370407
PR	13-OCT-2000	2000US-02393935
PR	13-OCT-2000	2000US-02393937
PR	13-OCT-2000	2000US-02393960
PR	20-OCT-2000	2000US-02412121
PR	20-OCT-2000	2000US-02411855
PR	20-OCT-2000	2000US-02411865
PR	20-OCT-2000	2000US-02411877
PR	20-OCT-2000	2000US-02418089
PR	20-OCT-2000	2000US-02418097
PR	20-OCT-2000	2000US-02418269
PR	01-NOV-2000	2000US-02446171
PR	08-NOV-2000	2000US-02464744
PR	08-NOV-2000	2000US-02464745
PR	08-NOV-2000	2000US-02464757
PR	08-NOV-2000	2000US-02464776
PR	08-NOV-2000	2000US-02464787
PR	08-NOV-2000	2000US-02465228
PR	08-NOV-2000	2000US-02465234
PR	08-NOV-2000	2000US-02465249
PR	08-NOV-2000	2000US-02465259
PR	08-NOV-2000	2000US-02465609
PR	08-NOV-2000	2000US-02466109
PR	08-NOV-2000	2000US-02466139
PR	08-NOV-2000	2000US-02466177
PR	17-NOV-2000	2000US-02492507
PR	17-NOV-2000	2000US-02492508
PR	17-NOV-2000	2000US-02492509
PR	17-NOV-2000	2000US-02492511
PR	17-NOV-2000	2000US-02492512
PR	17-NOV-2000	2000US-02492513
PR	17-NOV-2000	2000US-02492514
PR	17-NOV-2000	2000US-02492515

XX	17-NOV-2000;	2000US-0249215P.	
PR	17-NOV-2000;	2000US-0249216P.	
PR	17-NOV-2000;	2000US-0249217P.	
PR	17-NOV-2000;	2000US-0249218P.	
PR	17-NOV-2000;	2000US-0249219P.	
PR	17-NOV-2000;	2000US-0249220P.	
PR	17-NOV-2000;	2000US-0249221P.	
PR	17-NOV-2000;	2000US-0249222P.	
PR	17-NOV-2000;	2000US-0249223P.	
PR	17-NOV-2000;	2000US-0249224P.	
PR	17-NOV-2000;	2000US-0249225P.	
PR	17-NOV-2000;	2000US-0249226P.	
PR	17-NOV-2000;	2000US-0249227P.	
PR	17-NOV-2000;	2000US-0249228P.	
PR	17-NOV-2000;	2000US-0249229P.	
PR	17-NOV-2000;	2000US-0249230P.	
PR	17-NOV-2000;	2000US-0250160P.	
PR	01-DEC-2000;	2000US-0250391P.	
PR	05-DEC-2000;	2000US-0251030P.	
PR	05-DEC-2000;	2000US-0251988P.	
PR	05-DEC-2000;	2000US-0256719P.	
PR	06-DEC-2000;	2000US-0251479P.	
PR	08-DEC-2000;	2000US-0251856P.	
PR	08-DEC-2000;	2000US-0251868P.	
PR	08-DEC-2000;	2000US-0251869P.	
PR	08-DEC-2000;	2000US-0251899P.	
PR	08-DEC-2000;	2000US-0251990P.	
PR	11-DEC-2000;	2000US-0254097P.	
PR	05-JAN-2001;	2001US-0259678P.	
XX	(HUMA-)	HUMAN GENOME SCI INC.	
PA			
XX	Rosen CA,	Barash SC,	Ruben SM;
XX			
XX	WPI,	2001-483426/52.	
DR			
XX			
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,		
XX	useful for preventing, diagnosing and/or treating cancers and metastasis.		
PS	Disclosure; SEQ ID NO 29980; 3071pp + Sequence Listing; English.		
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)		
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic		
CC	activity, and can be used in gene therapy and vaccine production. (I)		
CC	proteins and polynucleotides may be used in the prevention, diagnosis and		
CC	treatment of diseases associated with inappropriate (I) expression. For		
CC	example, they may be used to treat disorders associated with decreased		
CC	expression by rectifying mutations or deletions in a patient's genome		
CC	that affect the activity of (I) by expressing inactive proteins or to		
CC	supplement the patients own production of (I). Additionally, (I)		
CC	polynucleotides may be used to produce the secreted (I), by inserting the		
CC	nucleic acids into a host cell and culturing the cell to express the		
CC	protein. (I) proteins and polynucleotides may be used to prevent,		
CC	diagnose and treat immune/haematopoietic-related diseases, especially		
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703		
CC	to AAK67694 represent human immune/haematopoietic antigen genomic		
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169		
CC	represent sequences used in the exemplification of the present invention		
XX			
XX	Sequence 4357 BP; 1312 A; 834 C; 950 G; 1261 T; 0 U; 0 Other;		
XX			
XX	Query Match	85.6%; Score 15.4; DB 4; Length 4357;	
XX	Best Local Similarity	94.1%; Pred. No. 3.7e+02;	
XX	Matches 16; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	2 GACTGTGATCTCTCCAT 18		
DB	3193 GACTGAGATCTCTCCAT 3209		
XX			
XX	RESULT 40		
XX	ABL19346/C		
XX	ABL19346 standard; DNA; 4532 BP.		
XX	ABL19346;		
XX			
XX	26-MAR-2002 (first entry)		
XX			

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9511.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical; gene; ds.  
 XX  
 XX Drosophila melanogaster.  
 XX  
 XX WO200171042-A2.  
 XX  
 XX 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 XX Claim 1; SEQ ID NO 9511; 21bp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
 CC AB572072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX  
 SQ Sequence 4532 BP; 1263 A; 1022 C; 975 G; 1272 T; 0 U; 0 Other;  
 Query Match 85.6%; Score 15.4; DB 4; Length 4532;  
 Best Local Similarity 94.1%; Pred. No. 3.7e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGGAATCCCTCA 17  
 Db 1649 CGACAGTGAATCTTCCA 1633  
 RESULT 41  
 ABL15658/c  
 ID ABL15658 standard; cDNA; 6446 BP.  
 XX  
 XX ABL15658;  
 AC  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41456.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical; gene; ss.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 XX  
 XX 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX

XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR  
 XX P-PSDB; AB571555.  
 DR  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 XX Claim 1; SEQ ID NO 41456; 21bp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
 CC AB572072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX  
 SQ Sequence 6446 BP; 1819 A; 1415 C; 1394 G; 1818 T; 0 U; 0 Other;  
 Query Match 85.6%; Score 15.4; DB 4; Length 6446;  
 Best Local Similarity 94.1%; Pred. No. 3.8e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGGAATCCCTCA 17  
 Db 3563 CGACAGTGAATCTTCCA 3547  
 RESULT 42  
 ADO47193  
 ID ADO47193 standard; DNA; 200000 BP.  
 XX  
 XX ADO47193;  
 AC  
 XX  
 XX 15-JUL-2004 (first entry)  
 DT  
 XX  
 DE DNA sequence of a human immunoglobulin heavy chain variable region.  
 XX  
 XX immunoglobulin; B cell; germline; region V; region D; region J;  
 KM recombination-promoting protein; recombination activating gene 1; RAG-1;  
 KM RAG-2; ss; terminal deoxynucleotidyltransferase; Tdt; V(D)J recombinase.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2004029249-A1.  
 XX  
 XX 08-APR-2004.  
 PD  
 XX  
 XX 30-SEP-2003; 2003WO-AU001286.  
 PF  
 XX  
 XX 30-SEP-2002; 2002US-0415024P.  
 PR  
 XX (CENT-) CENTENARY INST CANCER MEDICINE & CELL BI.  
 XX  
 XX Jolly C;  
 PI  
 XX  
 XX WPI; 2004-316126/29.  
 DR  
 XX New vector comprising one or more immunoglobulin regions selected from V,  
 PT D, and J regions, a 5' flanking region, and a 3' flanking region, useful  
 PT for reverting cell lines to a pro-B cell-like state or to a germline-like  
 PT state.  
 XX  
 XX Example 1; Page 234-283; 382bp; English.  
 PS  
 XX

CC The specification describes a method for the preparation of  
 CC immunoglobulins. The method uses a vector for reverting cell lines to a  
 CC pro-B cell-like state or to a germline-like state, by replacing one or  
 CC more of the immunoglobulin regions V, D, and J of the rearranged  
 CC immunoglobulin gene with the germ-line or unrearranged V, D or J regions.  
 CC The vector can comprise a polynucleotide sequence encoding a  
 CC recombination-promoting protein, selected from recombination activating  
 CC gene 1 (RAG-1), RAG-2, terminal deoxynucleotidyltransferase (Tdt). These  
 CC proteins collectively constitute a V(D)J recombinase. The method is  
 CC useful for the preparation of immunoglobulin heavy and/or light chains  
 CC and their fragments. The present sequence encodes a human immunoglobulin  
 CC heavy chain variable region, and is used as a template to construct  
 CC vectors for use in the method of the invention.

XX Sequence 20000 BP, 59140 A, 40391 C, 40764 G, 59705 T, 0 U, 0 Other;

Query Match 85.6%; Score 15.4; DB 12; Length 200000;  
 Best Local Similarity 94.1%; Pred. No. 5e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
 |||||  
 DB 107278 GACTGTGAATCCTCCAT 107294

RESULT 43

ACA61261 ID ACA61261 standard; DNA, 18 BP.

XX ACA61261;

DT 11-AUG-2003 (first entry)

XX Rat p53 antisense oligonucleotide.

XX Rat; ss; antisense; p53; bioactive agent delivery.

XX Rattus norvegicus.

PN US2003026831-A1.

XX 06-FEB-2003.

PF 22-APR-2002; 2002US-00131786.

PR 20-APR-2001; 2001US-0285337P.

PA (LAKK/) LAKKARAJU A.

PA (DUBI/) DUBINSKY J M.

PA (LOWN/) LOW N.

PA (RAHM/) RAHMAN Y.

PI Lakkaraju A, Dubinsky JM, Low N, Rahman Y;

PT WPI; 2003-439031/41.

PS Claim 21; Page 5; 56pp; English.

XX The invention relates to a pharmaceutical composition comprising an  
 CC anionic liposome, a bioactive agent, a cation and/or a buffer. The  
 CC composition is useful for delivering a bioactive agent to a target e.g.  
 CC cell of an organ (e.g. brain, central nervous system, peripheral nervous  
 CC system, liver, lung, larynx, bone marrow, spleen, kidney, lymphatic  
 CC system, hematopoietic system, gastric mucosa, small intestine, large  
 CC intestine, gall bladder, pancreas, salivary gland, testes, ovary, cervix,  
 CC uterus, muscle, skin, thyroid gland, parathyroid gland, adrenal gland,  
 CC connective tissue, chondroid tissue, blood vessel, macrophage, pleura,  
 CC placenta, a tumour or a growth), non-dividing cells, neuronal cells,  
 CC hippocampal neuronal cells, cell that expresses a receptor belonging to

CC the low-density lipoprotein (LDL) gene family, cell that possesses (LDL)  
 CC receptor-related protein (LRP) receptor, cell that possesses an endocytic  
 CC (LDL) lipoprotein receptor-related protein receptor, cell that possesses  
 CC a receptor that is expressed in mammalian central nervous system (CNS),  
 CC pluripotent cell, stem cell, embryonic stem cells and bone marrow  
 CC derived stem cells. The present sequence represents the rat p53 antisense  
 CC oligonucleotide

XX Sequence 18 BP; 4 A, 6 C, 3 G, 5 T, 0 U, 0 Other;

Query Match 83.3%; Score 15; DB 10; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CTGTGAATCCTCCAT 18  
 |||||  
 DB 1 CTGTGAATCCTCCAT 15

RESULT 44

AA297144/C ID AA297144 standard; cDNA, 1024 BP.

XX AA297144;

DT 18-APR-2000 (first entry)

XX Human prostate cancer differentially expressed gene #5.

XX Prostate cancer specific gene; cancer; tumour progression; diagnose;  
 XX hyperproliferative cell growth; prostatic disorder; treatment;  
 XX metastatic prostate cancer; benign prostate hyperplasia; BPH; ss.

XX Homo sapiens.

OS W0964594-A2.

PN 16-DEC-1999.

PD 10-JUN-1999; 99WO-US013181.

PF 11-JUN-1998; 98US-0088877P.

PR 09-JUN-1999; 99US-00284475.

XX (CHIR ) CHIRON CORP.

PA Astel JH, Carroll B, Endege WO, Ford DM, Monahan JB, Schlegel R;

PA Steimann KB, Zhang J;

DR WPI; 2000-116541/10.

PT New isolated prostate cancer specific nucleic acids, used to develop  
 PT products for the diagnosis and treatment of cancer.

PS Claim 12; Page 72; 212pp; English.

XX This sequence represents a prostate cancer specific nucleic acid  
 CC sequence. The invention relates to a method for diagnosing cancer, tumour  
 CC progression, hyperproliferative cell growth or accompanying biological  
 CC and physical manifestations. The method involves contacting the  
 CC biological sample with a probe that comprises a sequence capable of  
 CC hybridising to any of the 339 nucleotide sequences given in the  
 CC specification (see AA297140-297478) and detecting duplex formation. The  
 CC products and methods of the invention can be used for the diagnosis,  
 CC prognosis, and treatment of cancer, tumour progression,  
 CC hyperproliferative cell growth, and accompanying physical and biological  
 CC manifestations. They can be used particularly for prostatic disorders  
 CC such as metastatic prostate cancer, localised prostate cancer, or benign  
 CC prostate hyperplasia (BPH)

XX Sequence 1024 BP, 230 A, 217 C, 251 G, 191 T, 0 U, 135 Other;

Query Match 83.3%; Score 15; DB 3; Length 1024;

Best Local Similarity 93.8%; Pred. No. 5.3e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ACTGTAATCCTCCAT 18  
Db 574 ACTGTAATCCTCCAT 559

## RESULT 45

ABN66254/C  
ID ABN66254 standard; DNA; 1065 BP.

AC ABN66254;

DT 01-JUL-2002 (first entry)

DE Streptococcus polynucleotide SEQ ID NO 421.

KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;

OS Streptococcus agalactiae.

PN MO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

PA (CHIR-) CHIRON SPA.

PI (GENO-) INST GENOMIC RES.

PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C,

PI Tetteijn H;

DR WPI: 2002-352536/38.

PS P-PSDB; ABP25623.

XX Claim 7; Page 3197; 4525BP; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

CC the specification. The proteins have antibacterial and anti-inflammatory

CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and

CC antibodies that bind (1) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (1) are used to detect Streptococcus in a

CC biological sample. (1) is used to determine whether a compound binds to

CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (1) may be used to recombinantly produce (1) and may be

CC used in gene therapy. Antibodies to (1) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins

XX Sequence 1065 BP; 323 A; 211 C; 217 G; 314 T; 0 U; 0 Other;

XX Query Match 83.3%; Score 15; DB 6; Length 1065;

XX Best Local Similarity 100.0%; Pred. No. 5.3e+02;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CTGTGAATCCTCCAT 18  
Db 436 CTGTGAATCCTCCAT 422

## RESULT 46

ABN71527\_01  
Continuation (2 of 22) of ABN71527 from base 100001 (Streptococcus polynucleotide SEQ ID

WP Sequence split into 22 fragments LOCUS ABN71527 Accession ABN71527

WP Fragment Name

WP ABN71527\_00

WP ABN71527\_01

WP ABN71527\_02

WP ABN71527\_03

WP ABN71527\_04

WP ABN71527\_05

WP ABN71527\_06

WP ABN71527\_07

WP ABN71527\_08

WP ABN71527\_09

WP ABN71527\_10

WP ABN71527\_11

WP ABN71527\_12

WP ABN71527\_13

WP ABN71527\_14

WP ABN71527\_15

WP ABN71527\_16

WP ABN71527\_17

WP ABN71527\_18

WP ABN71527\_19

WP ABN71527\_20

WP ABN71527\_21

WP

WP

WP

WP

WP

WP

WP

WP

WP

WP

Query Match 83.3%; Score 15; DB 6; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CTGTGAATCCTCCAT 18  
Db 66795 CTGTGAATCCTCCAT 66809

## RESULT 47

ADL71108  
ID ADL71108 standard; DNA; 93 BP.

AC ADL71108;

DT 01-JUL-2004 (first entry)

DE p53 hairpin oligo p53-2.

XX ss; p53; hairpin; stem cell; short hairpin RNA; shRNA; anti-HIV;

XX cytotaxtic; cell therapy; HIV; cancer.

XX Unidentified.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

(COLD-) COLD SPRING HARBOR LAB.

PI Fridman J, Hannon GJ, Hemann M, Lowe SW, Paddison PJ, Zilfou J;

PI Dickins R, Carmell MA, Rosenquist TA;

DR WPI: 2004-316101/29.

XX Introducing into a subject stem cells with partial or complete loss of







XX ABL65765;  
AC  
XX 15-MAY-2002 (first entry)  
DT  
XX Lung cancer related gene sequence SEQ ID NO:4102.  
DE  
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;  
KW gene; ds.  
OS Homo sapiens.  
XX  
XX WO200194629-A2.  
PN  
XX 13-DEC-2001.  
PD  
XX 30-MAY-2001; 2001WO-US010838.  
PF  
XX 05-JUN-2000; 2000US-0209473P.  
PR  
XX 05-JUN-2000; 2000US-0209531P.  
PR  
XX 18-SEP-2000; 2000US-0231133P.  
PR  
XX 18-SEP-2000; 2000US-0233617P.  
PR  
XX 20-SEP-2000; 2000US-0234009P.  
PR  
XX 20-SEP-2000; 2000US-0234034P.  
PR  
XX 20-SEP-2000; 2000US-0234052P.  
PR  
XX 22-SEP-2000; 2000US-0234509P.  
PR  
XX 22-SEP-2000; 2000US-0234567P.  
PR  
XX 25-SEP-2000; 2000US-0234923P.  
PR  
XX 25-SEP-2000; 2000US-0234924P.  
PR  
XX 25-SEP-2000; 2000US-0235077P.  
PR  
XX 25-SEP-2000; 2000US-0235082P.  
PR  
XX 25-SEP-2000; 2000US-0235134P.  
PR  
XX 25-SEP-2000; 2000US-0235280P.  
PR  
XX 26-SEP-2000; 2000US-0235637P.  
PR  
XX 26-SEP-2000; 2000US-0235638P.  
PR  
XX 27-SEP-2000; 2000US-0235711P.  
PR  
XX 27-SEP-2000; 2000US-0235720P.  
PR  
XX 27-SEP-2000; 2000US-0235840P.  
PR  
XX 27-SEP-2000; 2000US-0235863P.  
PR  
XX 28-SEP-2000; 2000US-0236028P.  
PR  
XX 28-SEP-2000; 2000US-0236032P.  
PR  
XX 28-SEP-2000; 2000US-0236033P.  
PR  
XX 28-SEP-2000; 2000US-0236034P.  
PR  
XX 28-SEP-2000; 2000US-0236109P.  
PR  
XX 28-SEP-2000; 2000US-0236111P.  
PR  
XX 29-SEP-2000; 2000US-0236842P.  
PR  
XX 29-SEP-2000; 2000US-0236891P.  
PR  
XX 02-OCT-2000; 2000US-0237172P.  
PR  
XX 02-OCT-2000; 2000US-0237173P.  
PR  
XX 02-OCT-2000; 2000US-0237278P.  
PR  
XX 02-OCT-2000; 2000US-0237294P.  
PR  
XX 02-OCT-2000; 2000US-0237295P.  
PR  
XX 02-OCT-2000; 2000US-0237316P.  
PR  
XX 03-OCT-2000; 2000US-0237425P.  
PR  
XX 03-OCT-2000; 2000US-0237598P.  
PR  
XX 03-OCT-2000; 2000US-0237604P.  
PR  
XX 03-OCT-2000; 2000US-0237606P.  
PR  
XX 03-OCT-2000; 2000US-0237608P.  
PR  
XX 01-NOV-2000; 2000US-0244867P.  
PR  
XX 01-NOV-2000; 2000US-0245084P.  
XX  
XX (AVAL- ) AVALON PHARM.  
XX  
XX Young PB, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
XX WPI; 2002-188264/24.  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 4102; 44pp; English.  
PS  
XX The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms  
CC tumour  
XX  
XX Sequence 417 BP; 92 A; 130 C; 93 G; 101 T; 0 U; 1 Other;  
SQ  
XX  
XX Query Match 82.2%; Score 14.8; DB 6; Length 417;  
Best Local Similarity 88.9%; Pred. No. 6.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 CGACTGTGAATCCTCCAT 18  
Db 278 CGACTGTGAATCCACAT 295  
|||||  
|||||

Search completed: January 14, 2005, 16:13:50  
Job time : 359.263 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:01:39 ; Search time 913.263 Seconds  
(without alignments)  
932.058 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18

Sequence: 1 CGACTGCAATCCTCCAT 18

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database :

GenEmbl:\*  
1: gb\_ha:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_by:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	199	10	RATP53T02
2	18	100.0	513	10	AY009504
3	18	100.0	1019	10	AF209191
4	18	100.0	1176	6	RNU90328
5	18	100.0	1627	6	AX401815
6	18	100.0	1627	6	AX827834
7	18	100.0	1627	10	RNP53
8	18	100.0	9414	12	AY297714
9	18	100.0	165316	2	AC119115
10	18	100.0	192326	2	AC134317
11	18	100.0	1225077	2	AC136563
12	17	94.4	129489	10	AL713860
13	17	94.4	150816	2	AC119532
14	17	94.4	181075	10	CNS071P2
15	17	94.4	229363	2	AC094940
16	17	94.4	255706	2	AC094186
17	17	94.4	265074	2	AC103024
18	17	94.4	267779	2	AC112826
19	17	94.4	275838	2	AL627428

20	17	94.4	328868	2	AC118955
21	16.4	91.1	84	10	MMANT02
22	16.4	91.1	390	6	AX435258
23	16.4	91.1	836	3	AB038502
24	16.4	91.1	1096	10	MMU297973
25	16.4	91.1	1173	6	AR052877
26	16.4	91.1	1173	6	AR127895
27	16.4	91.1	1173	6	AR288189
28	16.4	91.1	1176	10	AY044188
29	16.4	91.1	1183	10	AF161020
30	16.4	91.1	1208	10	AF051368
31	16.4	91.1	1241	10	MUSP53B
32	16.4	91.1	1285	10	MUSP53A
33	16.4	91.1	1322	10	MUSP53C
34	16.4	91.1	1360	10	AB017815
35	16.4	91.1	1377	10	MMP53
36	16.4	91.1	1409	10	AF151353
37	16.4	91.1	1429	10	AB021961
38	16.4	91.1	1457	10	AB020317
39	16.4	91.1	1746	10	AY212017
40	16.4	91.1	1773	10	MMP53R
41	16.4	91.1	1782	10	BC005448
42	16.4	91.1	2130	10	K02110
43	16.4	91.1	2130	10	K02110
44	16.4	91.1	86146	2	AC023805
45	16.4	91.1	116898	10	AL731687
46	16.4	91.1	150937	9	AL445243
47	16.4	91.1	163447	2	AC102741
48	16.4	91.1	172568	2	BX324147
49	16.4	91.1	201602	2	AC074146
50	16.4	91.1	214262	2	AC148732
51	16.4	91.1	229482	2	AC074149
52	16.4	91.1	263036	2	AC105597
53	16.4	91.1	283735	2	AC098078
54	16.4	88.9	3890	14	RNMR81
55	16	88.9	104422	9	AC009656
56	16	88.9	139166	9	AC011750
57	16	88.9	140680	8	AF131112
58	16	88.9	146836	9	AC019317
59	16	88.9	151660	2	AC027450
60	16	88.9	154827	9	AL592221
61	16	88.9	170950	14	AY510475
62	16	88.9	174077	14	AF147806
63	16	88.9	177874	14	AF243438
64	16	88.9	180105	9	AC015756
65	16	88.9	182549	9	AC061939
66	16	88.9	184254	9	AC131935
67	16	88.9	197379	2	AC119278
68	16	88.9	206511	2	AC137798
69	16	88.9	211009	8	AF326781
70	16	88.9	212270	10	AC108811
71	16	88.9	238305	2	AC139741
72	15.4	85.6	296	9	AF027558
73	15.4	85.6	315	10	AF196265
74	15.4	85.6	319	10	AF196264
75	15.4	85.6	320	10	AF196260
76	15.4	85.6	325	10	AF196262
77	15.4	85.6	327	10	AF196263
78	15.4	85.6	347	10	AF196261
79	15.4	85.6	351	10	AF196255
80	15.4	85.6	352	10	AF196259
81	15.4	85.6	367	10	RNG101V3
82	15.4	85.6	367	10	RNMBEVA2
83	15.4	85.6	396	6	A42904
84	15.4	85.6	407	6	AR271438
85	15.4	85.6	436	10	MMU07878
86	15.4	85.6	420	10	RNU76836
87	15.4	85.6	525	6	AX867656
88	15.4	85.6	525	6	BD147718
89	15.4	85.6	527	6	AR413963
90	15.4	85.6	527	6	AX970797
91	15.4	85.6	527	6	BD109516
92	15.4	85.6	608	6	AX079605

AC118955 Rattus no  
X00876 Murine gene  
AX435258 Sequence  
AB038502 Caenorhab  
AU297973 Mus muscu  
AR052877 Sequence  
AR127895 Sequence  
AR288189 Sequence  
AY044188 Mus muscu  
AF161020 Mus muscu  
AF051368 Mus muscu  
M13873 Mouse p53 m  
M13877 Mouse p53 m  
AB017815 Mus muscu  
M13874 Mouse p53 m  
AB017815 Mus muscu  
X00741 Mouse mRNA  
AF151353 Mus muscu  
AB021961 Mus muscu  
AB020317 Mus muscu  
AY212017 Mus muscu  
X01237 Mouse mRNA  
BC005448 Mus muscu  
K02110 Mus musculus  
AC023805 Mus muscu  
AL731687 Mouse DNA  
AL445243 Human DNA  
AC102741 Mus muscu  
BX324147 Dantio rer  
AC074146 Mus muscu  
AC148732 Callithrix  
AC074149 Mus muscu  
AC105597 Rattus no  
AC098078 Rattus no  
J04357 Red clover  
AC009656 Homo sapi  
AC011750 Homo sapi  
AL133292 Arabidops  
AC019317 Homo sapi  
AC027450 Homo sapi  
AL592221 Human DNA  
AY510475 Gallid he  
AF147806 Gallid he  
AF243438 Gallid he  
AC015756 Homo sapi  
AC061939 Homo sapi  
AC131935 Homo sapi  
AC119278 Mus muscu  
AC137798 Homo sapi  
AF326781 Trifolium  
AC108811 Mus muscu  
AC139741 Gallus ga  
AF027558 Actus nan  
AF196265 Rattus no  
AF196264 Rattus no  
AF196260 Rattus no  
AF196262 Rattus no  
AF196263 Rattus no  
AF196261 Rattus no  
AF196255 Rattus no  
AF196259 Rattus no  
X08526 R. norvegicu  
X88527 R. norvegicu  
A42904 Sequence 36  
AR271438 Sequence  
U07878 Mus musculus  
U76836 Rattus norv  
AX867656 Sequence  
BD147718 Primer fo  
AR413963 Sequence  
AX970797 Sequence  
BD109516 EST and e  
AX079605 Sequence

C 93	15.4	85.6	826	8	BT004454	166	15.4	85.6	110000	2	AC140812_3	Continuation (4 of
94	15.4	85.6	833	8	HLJ224591	167	15.4	85.6	110000	2	LMFLCR36_08	Continuation (9 of
95	15.4	85.6	870	6	CQ780681	168	15.4	85.6	110586	2	AC136235	AC136235 Rattus no
96	15.4	85.6	870	6	CQ782219	C 169	15.4	85.6	113254	2	AL137009	AL137009 Human DNA
97	15.4	85.6	870	6	BD125350	C 170	15.4	85.6	114302	8	AL162413	AL162413 Human DNA
98	15.4	85.6	870	6	BD126928	C 171	15.4	85.6	118017	8	AC147013	AC147013 Medicago
C 99	15.4	85.6	886	10	MMU04624	C 172	15.4	85.6	119865	10	AC004407	AC004407 Mus muscu
100	15.4	85.6	994	8	HS1224549	C 173	15.4	85.6	126470	2	AL1390234	AL1390234 Human DNA
101	15.4	85.6	1007	8	HPR224547	C 174	15.4	85.6	132597	2	AC074014	AC074014 Homo sapi
102	15.4	85.6	1018	8	HBR224553	C 175	15.4	85.6	132715	5	BK465189	BK465189 Zebrafish
103	15.4	85.6	1018	8	HLO224551	C 176	15.4	85.6	132922	2	AC144322	AC144322 Macaca mu
104	15.4	85.6	1018	8	HP1224556	C 177	15.4	85.6	136625	2	AC120597	AC120597 Homo sapi
105	15.4	85.6	1019	8	HMA224548	C 178	15.4	85.6	138940	8	AC063972	AC063972 Genomic S
106	15.4	85.6	1019	8	HSO224557	C 179	15.4	85.6	142256	2	AC079297	AC079297 Homo sapi
107	15.4	85.6	1033	8	HVE224561	C 180	15.4	85.6	144335	2	AC136445	AC136445 Homo sapi
108	15.4	85.6	1034	8	HBT224564	C 181	15.4	85.6	146304	2	AC061967	AC061967 Homo sapi
C 109	15.4	85.6	1069	8	BT002894	C 182	15.4	85.6	148807	2	AC007028	AC007028 Homo sapi
C 110	15.4	85.6	1078	8	AY084550	C 183	15.4	85.6	150152	2	AL135502	AL135502 Homo sapi
C 111	15.4	85.6	1296	10	RRTCRAR	C 184	15.4	85.6	151500	2	AC008355	AC008355 Drosophill
112	15.4	85.6	1364	6	AX876492	C 185	15.4	85.6	153426	10	AC121949	AC121949 Mus muscu
113	15.4	85.6	1364	6	BD156157	C 186	15.4	85.6	153572	5	BK323992	BK323992 Zebrafish
114	15.4	85.6	1364	6	AK027524	C 187	15.4	85.6	154034	10	AL1772336	AL1772336 Mouse DNA
115	15.4	85.6	1478	9	BC041640	C 188	15.4	85.6	154060	2	AL1662885	AL1662885 Human DNA
116	15.4	85.6	1490	6	CQ783924	C 189	15.4	85.6	157051	9	AC011095	AC011095 Homo sapi
117	15.4	85.6	1490	6	BD127903	C 190	15.4	85.6	159094	5	AL929193	AL929193 Zebrafish
118	15.4	85.6	1490	6	AK075317	C 191	15.4	85.6	159274	9	AC112651	AC112651 Homo sapi
119	15.4	85.6	1496	9	BC012140	C 192	15.4	85.6	159830	10	AC145267	AC145267 Mus muscu
120	15.4	85.6	1517	9	BC004911	C 193	15.4	85.6	161566	2	AC136462	AC136462 Rattus no
121	15.4	85.6	1535	6	BD063240	C 194	15.4	85.6	162874	2	AC101892	AC101892 Mus muscu
C 122	15.4	85.6	1598	11	BV176766	C 195	15.4	85.6	164459	2	AC105335	AC105335 Mus muscu
C 123	15.4	85.6	1664	6	AX597496	C 196	15.4	85.6	166312	9	AC067807	AC067807 Homo sapi
C 124	15.4	85.6	1664	6	AF052149	C 197	15.4	85.6	166988	10	AC144552	AC144552 Mus muscu
C 125	15.4	85.6	1670	5	BC049435	C 198	15.4	85.6	167545	2	AC138799	AC138799 Homo sapi
C 126	15.4	85.6	1755	5	BD192217	C 199	15.4	85.6	167830	2	AC027768	AC027768 Homo sapi
C 127	15.4	85.6	1848	5	BC055232	C 200	15.4	85.6	169260	2	AC132232	AC132232 Mus muscu
C 128	15.4	85.6	1884	5	AB060223	C 201	15.4	85.6	170820	2	AC009057	AC009057 Homo sapi
C 129	15.4	85.6	1975	3	AY060981	C 202	15.4	85.6	171712	2	BX510660	BX510660 Danio rer
C 130	15.4	85.6	2003	3	BC009866	C 203	15.4	85.6	172194	9	AC092042	AC092042 Homo sapi
C 131	15.4	85.6	2045	3	BT001563	C 204	15.4	85.6	172608	10	AC146214	AC146214 Pan trogl
C 132	15.4	85.6	2148	6	COS98503	C 205	15.4	85.6	172608	10	AC140925	AC140925 Mus muscu
C 133	15.4	85.6	2214	6	AY625501	C 206	15.4	85.6	172616	9	AL157398	AL157398 Human DNA
C 134	15.4	85.6	2243	6	CQ592871	C 207	15.4	85.6	172914	2	AC015604	AC015604 Homo sapi
C 135	15.4	85.6	3336	6	BC048021	C 208	15.4	85.6	173069	2	AC129563	AC129563 Mus muscu
C 136	15.4	85.6	3703	6	CQ572543	C 209	15.4	85.6	173206	2	AC141248	AC141248 Homo sapi
C 137	15.4	85.6	4532	6	COS98502	C 210	15.4	85.6	173217	2	AC027785	AC027785 Homo sapi
C 138	15.4	85.6	5086	8	AY675586	C 211	15.4	85.6	173413	2	AC145309	AC145309 Homo sapi
C 139	15.4	85.6	6415	9	HSM806183	C 212	15.4	85.6	173765	2	AC090457	AC090457 Homo sapi
C 140	15.4	85.6	6446	6	CQ592870	C 213	15.4	85.6	173913	2	AC139181	AC139181 Pan trogl
C 141	15.4	85.6	34867	8	AP002030	C 214	15.4	85.6	174092	2	AC069129	AC069129 Homo sapi
C 142	15.4	85.6	35353	2	AC020055	C 215	15.4	85.6	174930	2	AC145889	AC145889 Homo sapi
C 143	15.4	85.6	39575	2	AC145683	C 216	15.4	85.6	175161	2	BX927396	BX927396 Danio rer
C 144	15.4	85.6	41131	2	AC091480_6	C 217	15.4	85.6	175524	2	AC012108	AC012108 Homo sapi
C 145	15.4	85.6	45558	9	AC105390_6	C 218	15.4	85.6	176653	2	AP001857	AP001857 Homo sapi
C 146	15.4	85.6	46643	10	AL173512	C 219	15.4	85.6	177433	2	AP001458	AP001458 Homo sapi
C 147	15.4	85.6	55520	9	HS004884	C 220	15.4	85.6	177990	2	AC090306	AC090306 Homo sapi
C 148	15.4	85.6	56171	2	AC14872	C 221	15.4	85.6	178438	2	AC149105	AC149105 Homo sapi
C 149	15.4	85.6	61223	2	AC142438	C 222	15.4	85.6	178500	2	AL135926	AL135926 Homo sapi
C 150	15.4	85.6	65449	2	AC105402	C 223	15.4	85.6	179632	2	AC102797	AC102797 Homo sapi
C 151	15.4	85.6	69619	2	AC103638	C 224	15.4	85.6	179711	2	AC147867	AC147867 Gopherus
C 152	15.4	85.6	81372	2	AC139056	C 225	15.4	85.6	180242	2	AC141310	AC141310 Homo sapi
C 153	15.4	85.6	84365	2	AC147867	C 226	15.4	85.6	180397	2	AC007035	AC007035 Homo sapi
C 154	15.4	85.6	87077	2	HUAC004097	C 227	15.4	85.6	181142	3	AC008346	AC008346 Drosophill
C 155	15.4	85.6	89934	8	ATAC009337	C 228	15.4	85.6	181380	10	AC115123	AC115123 Mus muscu
C 156	15.4	85.6	93063	2	AP000644	C 229	15.4	85.6	182147	2	AC062032	AC062032 Homo sapi
C 157	15.4	85.6	93683	2	AC138685	C 230	15.4	85.6	182196	2	AP002773	AP002773 Homo sapi
C 158	15.4	85.6	100099	10	AE008684_3	C 231	15.4	85.6	182952	2	CNS06C80	AC139154 Human chr
C 159	15.4	85.6	108720	8	AE046436	C 232	15.4	85.6	183412	2	AP001899	AP001899 Homo sapi
C 160	15.4	85.6	109274	8	AC138010	C 233	15.4	85.6	184164	2	AC091020	AC091020 Homo sapi
C 161	15.4	85.6	110000	2	AC096087_2	C 234	15.4	85.6	185147	10	AC125521	AC125521 Mus muscu
C 162	15.4	85.6	110000	2	AC195683	C 235	15.4	85.6	185155	2	BX957362	BX957362 Danio rer
C 163	15.4	85.6	110000	2	AC123419_2	C 236	15.4	85.6	185670	3	AC008097	AC008097 Drosophill
C 164	15.4	85.6	110000	2	AC138332_2	C 237	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 165	15.4	85.6	110000	2	AC140811_3	C 238	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl

239	15.4	85.6	187415	2	AC026352	AC026352 Homo sapi	312	15.4	85.6	246199	2	AC127759	AC127759 Rattus no
240	15.4	85.6	187562	2	AC140452	AC140452 Mus muscu	313	15.4	85.6	246731	2	AC097731	AC097731 Rattus no
241	15.4	85.6	187741	2	AC087681	AC087681 Homo sapi	314	15.4	85.6	248926	2	AC111553	AC111553 Rattus no
242	15.4	85.6	188197	2	AC104892	AC104892 Mus muscu	315	15.4	85.6	249433	2	AC128523	AC128523 Rattus no
243	15.4	85.6	188406	2	AC140658	AC140658 Homo sapi	316	15.4	85.6	250147	10	AC122191	AC122191 Mus muscu
244	15.4	85.6	188436	2	AC120068	AC120068 Rattus no	317	15.4	85.6	250498	2	AC118113	AC118113 Rattus no
245	15.4	85.6	188734	2	AC069443	AC069443 Homo sapi	318	15.4	85.6	250810	2	AC103486	AC103486 Rattus no
246	15.4	85.6	189549	10	AC122880	AC122880 Mus muscu	319	15.4	85.6	251426	2	AC111816	AC111816 Rattus no
247	15.4	85.6	189689	10	AC123847	AC123847 Mus muscu	320	15.4	85.6	252435	2	AC094150	AC094150 Rattus no
248	15.4	85.6	190597	10	AC140371	AC140371 Mus muscu	321	15.4	85.6	256038	2	AC135932	AC135932 Rattus no
249	15.4	85.6	191533	10	AC123041	AC123041 Mus muscu	322	15.4	85.6	256381	2	AC103510	AC103510 Homo sapi
250	15.4	85.6	191854	2	AC141402	AC141402 Homo sapi	323	15.4	85.6	258648	2	AC130010	AC130010 Rattus no
251	15.4	85.6	192607	2	AC139984	AC139984 Rattus no	324	15.4	85.6	259730	2	AC094880	AC094880 Rattus no
252	15.4	85.6	193142	2	AC113861	AC113861 Rattus no	325	15.4	85.6	262239	2	AC147863	AC147863 Gallus ga
253	15.4	85.6	193291	10	AC132871	AC132871 Mus muscu	326	15.4	85.6	266328	2	AC094326	AC094326 Rattus no
254	15.4	85.6	193377	2	AC135187	AC135187 Mus muscu	327	15.4	85.6	274928	2	AC114717	AC114717 Rattus no
255	15.4	85.6	193513	2	AC074372	AC074372 Homo sapi	328	15.4	85.6	274993	2	AC133403	AC133403 Rattus no
256	15.4	85.6	196249	10	AC092202	AC092202 Mus muscu	329	15.4	85.6	276063	2	AC114458	AC114458 Rattus no
257	15.4	85.6	196838	2	AC115196	AC115196 Rattus no	330	15.4	85.6	281529	2	AC111407	AC111407 Rattus no
258	15.4	85.6	198348	9	AC141283	AC141283 Homo sapi	331	15.4	85.6	281797	3	AB003794	AB003794 Drosoephil
259	15.4	85.6	198348	9	AC148025	AC148025 Homo sapi	332	15.4	85.6	282895	2	AC096056	AC096056 Rattus no
260	15.4	85.6	198992	10	AC124506	AC124506 Mus muscu	333	15.4	85.6	284938	2	AC104173	AC104173 Rattus no
261	15.4	85.6	199180	2	AC141078	AC141078 Homo sapi	334	15.4	85.6	285939	2	AC107599	AC107599 Rattus no
262	15.4	85.6	199868	9	AB019440	AB019440 Homo sapi	335	15.4	85.6	293854	2	AC126646	AC126646 Rattus no
263	15.4	85.6	200000	9	AB019440	AB019440 Homo sapi	336	15.4	85.6	293854	2	AC126646	AC126646 Rattus no
264	15.4	85.6	200039	2	AC133765	AC133765 Rattus no	337	15.4	85.6	299420	2	AC094449	AC094449 Rattus no
265	15.4	85.6	200859	2	AC118255	AC118255 Mus muscu	338	15.4	85.6	299749	3	AB003606	AB003606 Drosoephil
266	15.4	85.6	201214	2	AC124928	AC124928 Rattus no	339	15.4	85.6	299876	2	AC106196	AC106196 Rattus no
267	15.4	85.6	201641	9	AP002514	AP002514 Homo sapi	340	15.4	85.6	301660	2	AC112802	AC112802 Rattus no
268	15.4	85.6	202375	2	AC116494	AC116494 Mus muscu	341	15.4	85.6	304186	2	AC112438	AC112438 Rattus no
269	15.4	85.6	202505	2	AC090232	AC090232 Homo sapi	342	15.4	85.6	305730	2	AC114347	AC114347 Rattus no
270	15.4	85.6	202561	2	CR354554	CR354554 Danio rer	343	15.4	85.6	311823	2	AC095362	AC095362 Rattus no
271	15.4	85.6	203260	2	AC102839	AC102839 Mus muscu	344	15.4	85.6	313691	2	AC096152	AC096152 Rattus no
272	15.4	85.6	205174	9	AC011804	AC011804 Homo sapi	345	15.4	85.6	335400	2	AC131459	AC131459 Rattus no
273	15.4	85.6	205515	2	AC110637	AC110637 Rattus no	346	15.4	85.6	336001	2	AC098369	AC098369 Rattus no
274	15.4	85.6	205885	2	AC141611	AC141611 Pan trogl	347	15.4	85.6	471	11	BV045647	BV045647 S212P6650
275	15.4	85.6	206056	9	AC141453	AC141453 Homo sapi	348	15.4	85.6	570	11	BV075016	BV075016 S212P6650
276	15.4	85.6	206536	2	AC145310	AC145310 Homo sapi	349	15.4	85.6	669	11	BV020984	BV020984 S212P66022
277	15.4	85.6	206925	2	AC133345	AC133345 Rattus no	350	15.4	85.6	1024	6	BD228750	BD228750 Genes and
278	15.4	85.6	207322	2	AC144632	AC144632 Mus muscu	351	15.4	85.6	1024	6	AR243977	AR243977 Sequence
279	15.4	85.6	207444	2	AC121502	AC121502 Mus muscu	352	15.4	85.6	1066	6	CO643464	CO643464 Sequence
280	15.4	85.6	208406	2	AC133405	AC133405 Rattus no	353	15.4	85.6	1065	6	AX606553	AX606553 Sequence
281	15.4	85.6	208938	10	AC125395	AC125395 Mus muscu	354	15.4	85.6	1274	5	BC056576	BC056576 Danio rer
282	15.4	85.6	210843	10	AC114593	AC114593 Mus muscu	355	15.4	85.6	2387	9	HS118182	HS118182 Homo sapien
283	15.4	85.6	211018	2	AC140879	AC140879 Homo sapi	356	15.4	85.6	5167	10	RATTSRB	RATTSRB Rattus norv
284	15.4	85.6	211250	2	AC090702	AC090702 Homo sapi	357	15.4	85.6	22403	3	AB014199	AB014199 Streptoco
285	15.4	85.6	212345	2	BX890581	BX890581 Danio rer	358	15.4	85.6	22405	1	U70854	U70854 Caenorhabdi
286	15.4	85.6	214428	2	AC141672	AC141672 Mus muscu	359	15.4	85.6	35191	9	AC111196	AC111196 Homo sapi
287	15.4	85.6	214504	2	AC141622	AC141622 Homo sapi	360	15.4	85.6	40360	6	AX602198	AX602198 Sequence
288	15.4	85.6	214511	2	AC129256	AC129256 Rattus no	361	15.4	85.6	40502	9	BX072565	BX072565 Human DNA
289	15.4	85.6	215036	2	AC129359	AC129359 Rattus no	362	15.4	85.6	64120	2	AC014287	AC014287 Drosoephil
290	15.4	85.6	215213	10	AC111135	AC111135 Mus muscu	363	15.4	85.6	70515	10	AL806531	AL806531 Mouse DNA
291	15.4	85.6	218855	2	AC122871	AC122871 Mus muscu	364	15.4	85.6	79388	9	AL1391983	AL1391983 Human DNA
292	15.4	85.6	219412	10	AL772341	AL772341 Mouse DNA	365	15.4	85.6	87579	9	AC105289	AC105289 Homo sapi
293	15.4	85.6	220173	9	AC012183	AC012183 Homo sapi	366	15.4	85.6	90370	10	BX005084	BX005084 Mouse DNA
294	15.4	85.6	220595	2	AC128899	AC128899 Rattus no	367	15.4	85.6	96411	10	AL663114	AL663114 Mouse DNA
295	15.4	85.6	224026	10	AC095402	AC095402 Rattus no	368	15.4	85.6	106873	2	AC094556	AC094556 Homo sapi
296	15.4	85.6	224421	10	AC117633	AC117633 Mus muscu	369	15.4	85.6	119895	1	AP004310	AP004310 Synchrocy
297	15.4	85.6	224890	2	AC118036	AC118036 Rattus no	370	15.4	85.6	122208	2	AL583826	AL583826 Human DNA
298	15.4	85.6	229336	2	AC131843	AC131843 Rattus no	371	15.4	85.6	133085	3	AC073510	AC073510 Homo sapi
299	15.4	85.6	229893	2	AC130032	AC130032 Rattus no	372	15.4	85.6	143136	9	AC023677	AC023677 Drosoephil
300	15.4	85.6	231171	2	AC132695	AC132695 Rattus no	373	15.4	85.6	148309	2	AC146529	AC146529 Macaca mu
301	15.4	85.6	232508	2	AC094043	AC094043 Rattus no	374	15.4	85.6	153471	2	AC027076	AC027076 Homo sapi
302	15.4	85.6	233447	2	AC103192	AC103192 Rattus no	375	15.4	85.6	155595	2	AC011880	AC011880 Homo sapi
303	15.4	85.6	235183	2	AC079427	AC079427 Mus muscu	376	15.4	85.6	156842	2	BX119928	BX119928 Homo sapi
304	15.4	85.6	235452	2	AC128416	AC128416 Rattus no	377	15.4	85.6	158719	2	AC010008	AC010008 Drosoephil
305	15.4	85.6	236164	10	AC101205	AC101205 Mus muscu	378	15.4	85.6	163711	3	AC093436	AC093436 Drosoephil
306	15.4	85.6	239282	2	AC096176	AC096176 Rattus no	379	15.4	85.6	171810	2	AC128062	AC128062 Rattus no
307	15.4	85.6	239312	2	AC112572	AC112572 Rattus no	380	15.4	85.6	172303	9	AC018474	AC018474 Homo sapi
308	15.4	85.6	241331	2	AC137301	AC137301 Rattus no	381	15.4	85.6	173599	2	AC141388	AC141388 Rattus no
309	15.4	85.6	243032	2	AC129167	AC129167 Rattus no	382	15.4	85.6	178916	2	AP002874	AP002874 Homo sapi
310	15.4	85.6	245698	2	AC124841	AC124841 Rattus no	383	15.4	85.6	179443	2	AC079051	AC079051 Homo sapi
311	15.4	85.6	246021	10	AL663037	AL663037 Mouse DNA	384	15.4	85.6	179954	9	AC007271	AC007271 Homo sapi

385	15	83.3	182423	10	AC126942	Mus muscu	C 458	14.8	82.2	2489	6	AX702458	AX702458 Sequence
386	15	83.3	183101	9	AC021915	Homo sapi	C 459	14.8	82.2	2498	6	AX275969	AX275969 Sequence
387	15	83.3	186540	9	AC012478	Homo sapi	C 460	14.8	82.2	2542	3	AY118550	AY118550 Drosophill
388	15	83.3	186674	2	AC091544	Homo sapi	C 461	14.8	82.2	2557	6	C0613581	C0613581 Sequence
389	15	83.3	187234	2	AC114142	Rattus no	C 462	14.8	82.2	2832	6	C0577285	C0577285 Sequence
390	15	83.3	187517	5	AY519500	Gallus ga	C 463	14.8	82.2	2908	8	AY253257	AY253257 Dblptaxi
391	15	83.3	187691	10	AC083816	Mus muscu	C 464	14.8	82.2	2908	3	AF227210	AF227210 Drosophill
392	15	83.3	196661	10	AC119967	Mus muscu	C 465	14.8	82.2	4050	3	AK100406	AK100406 Oryza sat
393	15	83.3	205350	2	AC078946	Mus muscu	C 466	14.8	82.2	4388	6	C0604445	C0604445 Sequence
394	15	83.3	207161	2	AC141175	Rattus no	C 467	14.8	82.2	4681	6	C0613580	C0613580 Sequence
395	15	83.3	208600	9	AP002348	Homo sapi	C 468	14.8	82.2	4725	6	C0613466	C0613466 Sequence
396	15	83.3	212974	2	BX327371	Danio rer	C 469	14.8	82.2	5974	2	AC018252	AC018252 Drosophill
397	15	83.3	215127	10	AC113542	Mus muscu	C 470	14.8	82.2	9275	14	AB017037	AB017037 Himetobi
398	15	83.3	216449	2	AC051616	Mus muscu	C 471	14.8	82.2	9275	14	AB183472	AB183472 Himetobi
399	15	83.3	220674	2	BX548000	Danio rer	C 472	14.8	82.2	11039	1	AE007268	AE007268 Sinorhizo
400	15	83.3	221357	2	AP004246	Homo sapi	C 473	14.8	82.2	11279	8	OSAS35073	OSAS35073 Oryza sat
401	15	83.3	221941	9	AC091060	Homo sapi	C 474	14.8	82.2	12631	1	AE004252	AE004252 Vibrio ch
402	15	83.3	222808	2	AC095280	Rattus no	C 475	14.8	82.2	14093	1	AE000860	AE000860 Methanoba
403	15	83.3	224788	2	AP001905	Homo sapi	C 476	14.8	82.2	18637	3	CER825F01	CER825F01
404	15	83.3	226505	2	AC006899	Caenorhab	C 477	14.8	82.2	20809	3	CER044C1	CER044C1
405	15	83.3	235504	2	AC102729	Mus muscu	C 478	14.8	82.2	27365	2	AC015209	AC015209 Drosophill
406	15	83.3	237644	2	CR318653	Danio rer	C 479	14.8	82.2	30332	3	CER043C3	CER043C3
407	15	83.3	240050	1	SAG766844	Streptococ	C 480	14.8	82.2	38388	2	AC019820	AC019820 Drosophill
408	15	83.3	241173	2	AC126150	Rattus no	C 481	14.8	82.2	39216	3	CER02250	CER02250
409	15	83.3	241781	2	AC122087	Rattus no	C 482	14.8	82.2	42035	6	C0363774	C0363774 Sequence
410	15	83.3	243619	2	AC133838	Rattus no	C 483	14.8	82.2	42241	2	CER088B6	CER088B6
411	15	83.3	259619	2	AC108308	Rattus no	C 484	14.8	82.2	42589	2	AC124053	AC124053 Mus muscu
412	15	83.3	266775	2	AC129793	Rattus no	C 485	14.8	82.2	68663	2	AC116750	AC116750 Mus muscu
413	15	83.3	269788	10	AC125028	Mus muscu	C 486	14.8	82.2	69446	2	AC137105	AC137105 Mus muscu
414	15	83.3	272563	2	AC095469	Rattus no	C 487	14.8	82.2	74149	2	AC1238919_3	AC1238919_3
415	15	83.3	279448	8	AY661656	Sorghum b	C 488	14.8	82.2	85095	3	AC0004573	AC0004573 Drosophill
416	15	83.3	289192	2	AC139506	Homo sapi	C 489	14.8	82.2	89464	2	AC138543	AC138543 Magnapor
417	15	83.3	297981	3	AE003540	Drosophill	C 490	14.8	82.2	90544	2	AC136262	AC136262 Rattus no
418	15	83.3	349980	6	C0654010	Sequence	C 491	14.8	82.2	93889	2	AC105432	AC105432 Magnapor
419	15	83.3	349980	6	AX954527	Sequence	C 492	14.8	82.2	95992	2	AC138374	AC138374 Homo sapi
420	14.8	82.2	250	11	BTU94822ST	X98436 B.taurus an	C 493	14.8	82.2	96930	8	AC149464	AC149464 Populus b
421	14.8	82.2	350	6	C0685634	Sequence	C 494	14.8	82.2	97862	5	BX936370	BX936370 Zebrafish
422	14.8	82.2	417	6	AX333593	Sequence	C 495	14.8	82.2	100000	9	AP0000071	AP0000071 Homo sapi
423	14.8	82.2	417	6	AX334292	Sequence	C 496	14.8	82.2	102331	8	AC0007120	AC0007120 Arabidops
424	14.8	82.2	417	6	AX407434	Sequence	C 497	14.8	82.2	106073	9	AL136146	AL136146 Human DNA
425	14.8	82.2	482	14	AY636364	AY636364 HIV-1 93S	C 498	14.8	82.2	110000	1	AB017283_14	Continuation (15 of
426	14.8	82.2	527	5	CFCGRXS1	AY636364 HIV-1 93S	C 499	14.8	82.2	110000	2	AC091229_14	Continuation (5 of
427	14.8	82.2	635	11	BV026300	AY636364 HIV-1 93S	C 500	14.8	82.2	110000	2	AC091242_4	Continuation (5 of
428	14.8	82.2	636	14	AY636365	AY636365 HIV-1 99S							
429	14.8	82.2	687	14	HIVB8LTR	L28890 Human Immun							
430	14.8	82.2	714	14	AF239657	AF239657 HIV-1 iso							
431	14.8	82.2	716	14	AF096643	AF096643 HIV-1 iso							
432	14.8	82.2	763	14	AY221666	AY221666 HIV-1 iso							
433	14.8	82.2	791	6	CO431966	CO431966 Sequence							
434	14.8	82.2	796	14	AF62722	AF62722 HIV-1 iso							
435	14.8	82.2	824	14	AF62721	AF62721 HIV-1 iso							
436	14.8	82.2	928	6	CQ729867	CQ729867 Sequence							
437	14.8	82.2	930	9	AF173378	AF173378 Homo sapi							
438	14.8	82.2	966	6	AX275970	AX275970 Sequence							
439	14.8	82.2	1002	6	CO363925	CO363925 Sequence							
440	14.8	82.2	1016	9	AY303790	AY303790 Homo sapi							
441	14.8	82.2	1043	9	BC003013	BC003013 Homo sapi							
442	14.8	82.2	1043	9	BC006504	BC006504 Homo sapi							
443	14.8	82.2	1170	6	AX883719	AX883719 Sequence							
444	14.8	82.2	1170	6	BD160491	BD160491 Primer fo							
445	14.8	82.2	1170	9	AK024227	AK024227 Homo sapi							
446	14.8	82.2	1185	6	AX876719	AX876719 Sequence							
447	14.8	82.2	1185	6	BD156279	BD156279 Primer fo							
448	14.8	82.2	1185	6	AK027569	AK027569 Homo sapi							
449	14.8	82.2	1188	6	AX327430	AX327430 Sequence							
450	14.8	82.2	1447	10	RATP5EUD	L12046 Rattus norv							
451	14.8	82.2	1508	8	AF140552	AF140552 Lycopersi							
452	14.8	82.2	2040	4	OCRNAP53	X30592 O.cuniculus							
453	14.8	82.2	2137	3	AK064933	AK064933 Oryza sat							
454	14.8	82.2	2340	3	AY437499	AY437499 Tryptanso							
455	14.8	82.2	2368	3	AY634227	AY634227 Oikopleur							
456	14.8	82.2	2388	6	CO604446	CO604446 Sequence							
457	14.8	82.2	2488	3	DM047619	DM047619 Drosophila							

LOCUS	RATP53T502/c	199 bp	DNA	linear	ROD 14-JUL-1993
DEFINITION	Rattus norvegicus tumor suppressor (p53) gene, exon 2.				
ACCESSION	L07904	L07781			
VERSION	L07904.1	GI:205944			
KEYWORDS	tumor suppressor.				
SOURCE	2 of 8				
ORGANISM	Rattus norvegicus (Norway rat)				
REFERENCE	1 (bases 1 to 199)				
AUTHORS	Hulla,J.E. and Schneider,R.P.				
TITLE	Structure of the rat p53 tumor suppressor gene				
JOURNAL	Nucleic Acids Res. 21 (3), 713-717 (1993)				
MEDLINE	93181268				
PubMed	8441680				
COMMENT	Original source text: Rattus norvegicus (strain Sprague-Dawley) DNA.				
FEATURES	From EMBL	entry V00947	dated 23-OCT-1992.		
source	1..199	Location/Qualifiers			
	/organism="Rattus norvegicus"				



intron

exon

ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 199;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 CGACTGTGAATCCTCCAT 50

Qy 1 CGACTGTGAATCCTCCAT 18  
|||||  
|||||

RESULT 2  
AY009504/c 513 bp mRNA linear ROD 28-NOV-2001  
LOCUS Rattus norvegicus mutant p53 (p53) mRNA, complete cds.  
ACCESSION AY009504  
VERSION AY009504.1 GI:11528485  
KEYWORDS  
SOURCE  
ORGANISM Rattus norvegicus (Norway rat)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 513)  
Geutskens, S.B., Van den Mollenberg, D.J.M., Van der Eb, M.M., Van  
Ormondt, H., Jochemsen, A.G. and Hoeber, R.C.  
Characterisation of the p53 gene in the rat CC531 colon carcinoma  
2 (bases 1 to 513)  
Gene Ther. Mol. Biol. 5, 81-86 (2000)  
Hoeber, R.C., Van der Eb, M.M. and Geutskens, S.B.  
Direct Submission  
Submitted (13-OCT-2000) Dept. of Molecular Cell Biology, Molecular  
Virology Laboratory, Massenaarsweg 72, Leiden, Zuid Holland 2333  
AL, Netherlands  
Location/Qualifiers  
1. 513  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="WAG/Rij"  
/db\_xref="taxon:10116"  
/tissue\_type="CC531 colon carcinoma"  
1. 513  
/gene="p53"  
1. 513  
/gene="p53"  
/note="entire DNA binding domain is deleted"  
/codon\_start=1  
/product="mutant p53"  
/protein\_id="AA035178.1"  
/db\_xref="GI:11528485"  
/translation="MEDSDMSIFLPLSQETSCIMKLLPPDILPTTAGSPNSME  
DLFLPDVALELGGPEALQVAPAAQEPGTEAPVAPASATPPLSSVPSQKTYQ  
GNFLTKINGRRFEMFRLNALDELKQARAASGSDSRHSSYPKTKGQSTSRHKP  
MKIKVGPSD"

ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 513;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18  
|||||  
|||||

Db 18 CGACTGTGAATCCTCCAT 1

RESULT 3  
AF209191/c 1019 bp mRNA linear ROD 01-DEC-2000  
LOCUS Rattus norvegicus p53 alternative splice isoform p35/HAS mRNA,  
complete cds.  
ACCESSION AF209191  
VERSION AF209191.1 GI:11494109  
KEYWORDS  
SOURCE  
ORGANISM Rattus norvegicus (Norway rat)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1019)  
Dell'Acqua, G., Mann, M.J., Mu, H., Griese, D.P., Ehsan, A. and Dzau, V.  
Unpublished  
2 (bases 1 to 1019)  
Dell'Acqua, G., Mann, M.J., Mu, H., Griese, D.P., Ehsan, A. and Dzau, V.  
Direct Submission  
Submitted (29-NOV-1999) Medicine, Brigham and Women's  
Hospital/Harvard Medical School, 75 Francis St, Boston, MA 02215,  
USA  
Location/Qualifiers  
1. 1019  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
24. 839  
/note="frameshift results in premature termination"  
/codon\_start=1  
/product="p53 alternative splice isoform p35/HAS"  
/protein\_id="AA035765.1"  
/db\_xref="GI:11494110"  
/translation="MEDSDMSIFLPLSQETSCIMKLLPPDILPTTAGSPNSME  
DLFLPDVALELGGPEALQVAPAAQEPGTEAPVAPASATPPLSSVPSQKTYQ  
GNVGFHGFLOSQTSKSVMTYSLTKLFCOLAKTQPVOLMTSTPGTRVAMAI  
YKSGQHTVEVRRCPHERCSGDGLAPQGLIVVEGNPYAEYDDQTFEHSVAVPY  
EPPVGSDDYTHHKVMCNSSCMGMRRPILITITLEDSSRVSASRCSES"

ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18  
|||||  
|||||

Db 41 CGACTGTGAATCCTCCAT 24

RESULT 4  
RNU90328/c 1176 bp mRNA linear ROD 14-OCT-1997  
LOCUS Rattus norvegicus mutant p53 mRNA, complete cds.  
DEFINITION RNU90328  
ACCESSION U90328  
VERSION U90328.1 GI:1938364  
KEYWORDS  
SOURCE  
ORGANISM Rattus norvegicus (Norway rat)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1176)  
Mathupala, S.P., Heese, C. and Pedersen, P.L.  
Glucose catabolism in cancer cells. The type II hexokinase promoter  
contains functionally active response elements for the tumor  
suppressor p53  
Rattus.  
1 (bases 1 to 1176)  
Mathupala, S.P., Heese, C. and Pedersen, P.L.  
Glucose catabolism in cancer cells. The type II hexokinase promoter  
contains functionally active response elements for the tumor  
suppressor p53  
J. Biol. Chem. 272 (36), 22776-22780 (1997)  
MEDLINE 97426425

PUBMED 9278438  
REFERENCE 2 (bases 1 to 1176)  
AUTHORS Mathupala,S.P.  
TITLE Direct Submission  
JOURNAL Submitted (21-FEB-1997) Biological Chemistry, Johns Hopkins  
University School of Medicine, 725 N. Wolfe Street, Baltimore, MD  
21205, USA

FEATURES  
source  
1..1176  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/cell\_line="AS-30D"  
/cell\_type="ascites hepatoma"  
1..1176  
/function="enhances transcription of type II hexokinase"  
/codon\_start=1  
/product="mutant p53"  
/protein\_id="AAB80959.1"  
/db\_xref="GI:1938365"  
/translation="MEDSOSMSTELPLSOETPSCIMKLLPPDILPTATGPNSE  
DLRPPVVAELLEGPERALQYSABAQEPGTAPAPAPASATPWPSSSVPSKTIQ  
SNIFHLGFLQSGTAKSVCTYISLNLFCQLAKTCPOVLWVTSTPPGTRVAPAAI  
YKKSQNMTEVVRRCPHNRCSDGDLAPPOHLIRGNPYAEYLDROTFRHSVVVY  
EPPEVGSDDYTIHYKMCNSGCMGNRRPILITLIGDSSGNLIGRDSFEVRCAP  
GRDRTEENFRKKEHCPELPGSAKRALPTSTSSPOOKKPLDEYFTLKIGRE  
RFEMFREINALELKDARAAESGDSRAHSYPTKKGQSTSRHKMKIKVGPDDSD"

ORIGIN  
Query Match 100.0%; Score 18; DB 10; Length 1176;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 18 CGACTGTGAATCCTCCAT 1

RESULT 5  
AX401815/c 1627 bp DNA linear PAT 06-JUN-2002  
LOCUS AX401815  
DEFINITION Sequence 1491 from Patent WO0210453.  
ACCESSION AX401815  
VERSION AX401815.1 GI:21337995  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1  
AUTHORS Mendrick,D., Porter,M.W., Johnson,K.R., Castle,A.L. and  
Elaishoff,M.R.  
TITLE Molecular toxicology modeling  
JOURNAL Patent: WO 0210453-A 1491 07-FEB-2002;  
Gene Logic, Inc. (US)  
FEATURES  
source  
1..1627  
/organism="Rattus norvegicus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10116"  
/note="EMBL/GenBank Accession No. X13058"

ORIGIN  
Query Match 100.0%; Score 18; DB 6; Length 1627;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 41 CGACTGTGAATCCTCCAT 24

RESULT 6  
AX827834/c 1627 bp DNA linear PAT 12-DEC-2003  
LOCUS AX827834  
DEFINITION Sequence 568 from Patent EP1344834.  
ACCESSION AX827834  
VERSION AX827834.1 GI:39838022  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1  
AUTHORS Boess,F., Suter-Dick,L. and Wolf,D.  
TITLE Methods for the toxicity prediction of a compound  
JOURNAL Patent: EP 1344834-A 568 17-SEP-2003;  
F. HOFMANN-LA ROCHE AG (CH)  
FEATURES  
source  
1..1627  
/organism="Rattus norvegicus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10116"

ORIGIN  
Query Match 100.0%; Score 18; DB 6; Length 1627;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 41 CGACTGTGAATCCTCCAT 24

RESULT 7  
RNP53/c 1627 bp mRNA linear ROD 12-SEP-1993  
LOCUS RNP53  
DEFINITION Rat mRNA for nuclear oncoprotein p53.  
ACCESSION X13058  
VERSION X13058.1 GI:56828  
KEYWORDS oncoprotein p53.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 1627)  
AUTHORS Soussi,T., Caron de Fromental,C., Breugnot,C. and May,B.  
TITLE Nucleotide sequence of a cDNA encoding the rat p53 nuclear  
JOURNAL Nucleic Acids Res. 16 (23), 11384 (1988)  
MEDLINE 89083585  
PUBMED 3060862  
REFERENCE 2 (bases 1 to 1627)  
AUTHORS Soussi,T.  
TITLE Direct Submission  
JOURNAL Submitted (26-SEP-1988) Soussi T., Universite Pierre et Marie  
Curie, Unite d'Oncologie Molculaire, IRSC - CNRS, BP 08 94802  
Villejuif, France  
COMMENT  
data kindly reviewed (09-Feb-1989) by Soussi T.  
FEATURES  
source  
1..1627  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/cell\_line="Py T21"  
24..1199  
/note="unnamed protein product; nuclear protein p53 (AA 1  
-391)"  
/codon\_start=1  
/protein\_id="CAA31457.1"  
/db\_xref="GI:56829"  
/db\_xref="GOA:P10361"

/db\_xref="Swiss-Prot:P10361"  
 /translation="MEDSODMSIELPLSOFETSCMLKPLPPDILLPTTANSPNSME  
 DLPLPOVAELLEGREBALOVSAPAOBGTAPAPVAPASATPPLSSVPSOCTYO  
 GNYGPHLFLPSGRTKSWCTYSISLTKLFCOLATCTCVQMTSTPPPGTRFMAL  
 YKSSQHMTEVVRCPHRCRSCSDGLAPQHLIRFENPVAEYLDROTFRRSVVAPY  
 EPPGVSDYVTHHYKMCNCSGCMGNRRPILITITLSDSGNLLGRDSFEVAVAC  
 GRDRTEENFRKKEHCELPFGSAKRALPSTSSSQCKKGLDGGFTLKIGRE  
 RPEMRREINELKMDADAESGDSRAHSSVPKTKKGSIRHKKPMIKYKVPDSD"

## ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 1627;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGACTGTGAATCTCCAT 18  
 Db 41 CGACTGTGAATCTCCAT 24

## RESULT 8

AY297714 9414 bp DNA circular SYN 10-JUN-2003  
 DEFINITION Yeast truncation assay backbone vector pLSK870, complete sequence.  
 ACCESSION AY297714  
 VERSION AY297714.1 GI:31580803

## KEYWORDS

## SOURCE

Yeast truncation assay backbone vector pLSK870  
 Yeast truncation assay backbone vector pLSK870

## REFERENCE

## AUTHORS

1 (bases 1 to 9414)  
 Zan, Y., Haag, J.D., Chen, K.-S., Shepel, L.A., Wightington, D.,  
 Wang, Y.-R., Hu, R., Lopez-Guajardo, C.C., Brose, H.L., Porter, K.I.,  
 Leonard, R.A., Hitt, A.A., Schommer, S.L., Eliebede, A.F. and  
 Gould, M.N.

## TITLE

Production of knockout rats using ENT mutagenesis and a yeast-based  
 screening assay

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## TITLE

## JOURNAL

## FEATURES

## SOURCE

## misc\_feature

## promoter

## misc\_feature

## CDS

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

misc\_feature 4305..4312  
 /note="NotI cloning site"  
 misc\_feature 4315..6027  
 /note="encodes yeast AD22 reporter gene"

## ORIGIN

Query Match 100.0%; Score 18; DB 12; Length 9414;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGACTGTGAATCTCCAT 18  
 Db 4230 CGACTGTGAATCTCCAT 4273

## RESULT 9

AC119115 165316 bp DNA linear HTG 19-NOV-2002  
 DEFINITION Rattus norvegicus clone CH230-320N23, WORKING DRAFT SEQUENCE, 5  
 unordered pieces.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

## AUTHORS

## REFERENCE

Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstock, G. and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 165316)  
 Worley, K.C.  
 Direct Submission  
 Submitted (25-APR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 165316)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 19, 2002 this sequence version replaced gi:23616728.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GUXG  
 Center clone name: CH230-320N23  
 ----- Summary Statistics  
 Assembly program: Phrap; Version 0.990329  
 Consensus quality: 155971 bases at least Q40  
 Consensus quality: 157321 bases at least Q30  
 Consensus quality: 158253 bases at least Q20  
 Estimated insert size: 159662; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 159511: contig of 159511 bp in length  
 159512 159611: gap of unknown length  
 159612 160669: contig of 1058 bp in length  
 160670 160769: gap of unknown length  
 160770 161864: contig of 1095 bp in length  
 161865 161964: gap of unknown length  
 161965 163701: contig of 1737 bp in length  
 163702 163801: gap of unknown length  
 163802 165316: contig of 1515 bp in length.  
 Location/Qualifiers  
 1. 165316  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-320N23"

misc\_feature  
 1. 2103  
 /note="wgs\_end\_extension  
 clone\_end:17"  
 8259. 9143  
 /note="clone\_boundary  
 clone\_end:17  
 site:  
 end\_sequence:BZ171094"  
 46341. 159014  
 /note="clone\_boundary  
 clone\_end:5p6  
 site:  
 end\_sequence:BZ171096"

misc\_feature  
 end\_sequence:BZ171094"  
 46341. 159014  
 /note="clone\_boundary  
 clone\_end:5p6  
 site:  
 end\_sequence:BZ171096"

ORIGIN  
 Query Match 100.0%; Score 18; DB 2; Length 165316;  
 Best local similarity 100.0%; Pred. No. 17;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CGACTGGAATCCCTCAT 18  
 Db 14541 CGACTGGAATCCCTCAT 14524

RESULT 10  
 AC134317  
 LOCUS  
 DEFINITION  
 Rattus norvegicus clone RP31-421B16 strain Brown Norway, WORKING  
 DRAFT SEQUENCE, 12 ordered pieces.  
 AC134317  
 VERSION  
 AC134317.3 GI:31442441  
 HTG: HTGS PHASE2; HTGS DRAFT.  
 KEYWORDS  
 Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

1 (bases 1 to 192326)  
 Antonellis, A., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariga, K.,  
 Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,  
 Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,  
 Hurler, B., Idol, J.R., Karlins, E., Kwong, P., Laric, P., Lee-Lin, S.-O.,  
 Legaspi, R., Maduro, Q.L., Maduro, V.B., Marulles, E.H., Mastello, C.,  
 Maskeri, B., McDowell, J., Pagnitigan, C., Pearson, R., Portnoy, M.E.,  
 Prasad, A., Reddix-Dugue, N., Schandier, K., Schneider, M.G., Shan, K.,  
 Simon, C., Stancirip, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,  
 Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 192326)  
 Green, E.D.  
 Direct Submission  
 Submitted (25-SEP-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gailthersburg, MD 20877, USA  
 3 (bases 1 to 192326)  
 Green, E.D.  
 Direct Submission  
 Submitted (06-JUN-2003) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gailthersburg, MD 20877, USA  
 On Jun 6, 2003 this sequence version replaced gi:2753661.

----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: http://www.nisc.nih.gov  
 Contact: nisc\_zoo@ngi.nih.gov  
 ----- Project Information  
 Center project name: dce  
 Center clone name: 421B16

The sequence data in this record represents an 'enhanced'  
 version of a Phase 2 submission. Specifically, the indicated  
 order and orientation of each sequence contig has been

1 (buses sent to 225077)

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismato, K., Blair, J., Blankenbush, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burkell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessari, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Derick, D., Delgado, O., Denison, S., Berry, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Tocchi, S., Dunn, A., Dublin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falset, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabriel, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guenara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koval, C., Kowis, C., Kraft, C. L., Lebow, H., Leyvan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Lottesged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Mlilao, V., Mlile, A., Mliler, G., Minda, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Pasernak, S., Paul, H., Perez, A., Perez, L., Pfamko, C., Plopper, F., Poidexter, A., Popovic, D., Prims, E., Pu, J., Pua, M., Quito, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ritz, S. J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X., Z., Sorrelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Vals, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

Unpublished  
2 (bases 1 to 225077)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (06-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 225077)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:25188706.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Center project name: GHDO  
Center clone name: CH230-46E21  
----- Summary Statistics  
Assembly program: Atlas 3.0  
Consensus quality: 21459 bases at least Q40  
Consensus quality: 216978 bases at least Q30  
Consensus quality: 218480 bases at least Q20  
Estimated insert size: 225046; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)).  
\* NOTE: This is a working draft sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES

source

1 225077: contig of 225077 bp in length.  
Location/Qualifiers  
1..225077  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-46E21"  
1..1402  
/note="wgs\_contig"

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 225077;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTAATCTTCAT 18  
|||||  
Db 93482 CGACTGTAATCTTCAT 93465

RESULT 12  
AL713860  
LOCUS  
DEFINITION  
Mouse DNA sequence from clone RP23-446K18 on chromosome 11,  
complete sequence.  
ACCESSION  
AL713860  
VERSION  
AL713860.12 GI:27899576  
KEYWORDS  
HTG.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 129489)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (24-JAN-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
On Jan 24, 2003 this sequence version replaced gi:21217846.  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; SW, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)

-----  
Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
RP23-446K18 is from the RP21-23 Mouse BAC Library constructed by the group of Pieter de Jong.  
For further details see <http://www.choxi.org/bacpac/home.htm>  
VECTOR: pBAC3.6  
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
Location/Qualifiers

FEATURES



```

/note="clone boundary
clone_end:sp6
site:
end_sequence:BT174726"
misc_feature
/clone="clone_boundary
clone_end:t7
site:
end_sequence:BT174725"
misc_feature
/clone="wgs_end_extension
clone_end:t7"

ORIGIN
Query Match          94.4%; Score 17; DB 2; Length 150816;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
|||||
DB 81343 GACTGTGAATCCTCCAT 81327

RESULT 14
CNS07YP2 181075 bp DNA linear ROD 23-MAR-2002
LOCUS Mus musculus chromosome 11 region in the Cm locus area
DEFINITION (D1Mit37-Scy46) clone 330D23 of library Caltech CITB-BAC from
Chromosome 11 of Mus musculus (mouse).
ACCESSION AT1713885
VERSION 425A1-SP6; DDK syndrome; Ovum mutant; Scy41.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 181075)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
COMMENT F. Bahiet C., Baldacci, P.
Unité de Biologie du Développement, CNRS URA 1960, Institut
Pasteur, 25 rue du Dr Roux, 75724 Paris cedex 15, France.
FEATURES
source
1. 181075
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="330D23"
/clone_id="Caltech CITB-BAC"

ORIGIN
Query Match          94.4%; Score 17; DB 10; Length 181075;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
|||||
DB 57477 GACTGTGAATCCTCCAT 57493

RESULT 15
AC094940 229363 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-6E9, *** SEQUENCING IN PROGRESS ***
DEFINITION 5 unordered pieces.
ACCESSION AC094940

```

```

VERSION
AC094940.5 GI:30466994
KEYWORDS HTG, HTGS_PHASE1, HTGS_DRAFT, HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 229363)
Wuzny,D.,Wartle, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Bacc,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davis,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Dublin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falle,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gumaratne,P., Healand,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louieged,H., Lozada,R.J., Lu,X., Ma,J.,
Maneshwari,W., Mahindartne,M., Mahmoud,M., Mallory,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Muniesa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokelimen,O., Okunolu,G., Olarinmusaogun,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plappert,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.,
Plazmo,B., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Rally,B., Reilly,M., Ren,Y., Reuter,M., Richard,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,M., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,S., Sitter,C.D., Smay,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soes,J.,
Steinle,M., Strong,R., Sutton,A., Swack,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valae,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,K., White,F.,
Williams,G., Willison,R., Wleczky,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,Y.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,X., Zhou,X., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 229363)
Worley,K.C.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229363)
Rat Genome Sequencing Consortium.
REFERENCE Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 9, 2003 this sequence version replaced gi:2271362.
The sequence in this assembly is a combination of BAC based reads

```



and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

#### Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GBTU

Center clone name: CH230-6E9

Summary Statistics

Assembly program: Atlas;

Consensus quality: 208118 bases at least Q40

Consensus quality: 212027 bases at least Q30

Consensus quality: 214790 bases at least Q20

Estimated insert size: 226740; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbankdraft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 131104: contig of 131104 bp in length  
 \* 131105 131204: gap of unknown length  
 \* 131205 223436: contig of 92232 bp in length  
 \* 223437 223536: gap of unknown length  
 \* 223537 224657: contig of 1121 bp in length  
 \* 224658 224757: gap of unknown length  
 \* 224758 227566: contig of 2809 bp in length  
 \* 227567 227666: gap of unknown length  
 \* 227667 229363: contig of 1697 bp in length.

#### FEATURES

source

1. 229363  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-6E9"

#### ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 229363;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GACTGTGATCCTCCAT 18  
 |||||  
 Db 223332 GACTGTGATCCTCCAT 223348

RESULT 16  
 AC094186/c 255706 bp DNA linear HTG 09-MAY-2003  
 LOCUS Rattus norvegicus clone CH230-2K14, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 DEFINITION 2 unordered pieces.  
 AC094186  
 VERSION AC094186.6 GI:30467744  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus

#### REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 255706)

Murphy, D., Marle, Mezker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalbechi, V., Ayagi, A., Ayodeji, M., Bacc, B., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bielawski, K., Blair, D., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, S., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., L., Davis, C., Day, Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Dudin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, B., Geier, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Ledow, H., Levan, D., Lewis, L., Li, Z., Liu, J., Liu, Y., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshwaha, L., Lolliseed, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, B., Mangum, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, S., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackemeke, O., Okunomi, G., Olariu, G., Olariu, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldexter, A., Popovic, D., Primus, B., Pu, L., L., Puzo, M., Quiroz, J., Rachin, B., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shastman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slison, I., Slicer, C. D., Snaj, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Swalek, A., Taber, J., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Welser, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

#### TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

#### REFERENCE

AUTHORS

TITLE

JOURNAL

#### COMMENT

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 255706)  
 Direct Submission  
 Rat Genome Sequencing Consortium.  
 On May 9, 2003 this sequence version replaced gi:24942737.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

#### ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GADZ

Center clone name: CH230-2K14

Summary Statistics

Assembly program: Atlas;

Consensus quality: 244860 bases at least Q40

Consensus quality: 247754 bases at least Q30

Consensus quality: 249340 bases at least Q20

Estimated insert size: 268126; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drafft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drafft_data.html))  
 \* NOTE: This sequence may represent more than one clone.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved

#### FEATURES

##### source

1..255706

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-2K14"

38438..43555

/note="wgs contig"

43606..46302

/note="wgs contig"

##### misc\_feature

##### misc\_feature

##### misc\_feature

##### misc\_feature

##### ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 255706;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

##### QY

##### Db

##### 56167

##### GACTGTGAATCTCCAT 18

##### AC103024

##### LOCUS

##### DEFINITION

##### AC103024

##### AC103024

##### AC103024.5

##### HTGS\_PHA5R1

##### HTGS\_DNAFT

##### HTGS\_FULITOP

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

##### Rattus norvegicus

#### COMMENT

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

Allen, C., Allen, H., Alstrooks, S., Amin, A., Anguiano, D., Anyalebehi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, B., Barnstead, M., Benham, F., Bialski, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessar, H., Chen, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabist, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladik, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louieged, H., Lozano, R.J., Lu, X., Ma, U., Maheshwari, M., Mahindartne, N., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nuckelmeier, O., Okunolu, G., Olumunsegun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodery, T., Rojas, A., Rose, M., Rose, R., Ruiz, J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, B., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valae, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiczek, R., Woodem, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

#### JOURNAL

#### AUTHORS

#### TITLE

shotgun sequence only contigs will be indicated in the feature table.

# ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

# -----Project Information

Center project name: GIY

Center clone name: CH230-19984

# ----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 243702 bases at least Q40

Consensus quality: 245482 bases at least Q30

Estimated insert size: 255715; sum-of-contigs estimation

Quality coverage: 7x in Q30 bases; sum-of-contigs estimation

-----

\* NOTE: Rattified insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))

\* NOTE: This sequence may represent more than one clone.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 3 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 262317: contig of 262317 bp in length

\* 262318 262417: gap of unknown length

\* 262418 263766: contig of 1349 bp in length

\* 263767 263866: gap of unknown length

\* 263867 265074: contig of 1208 bp in length.

\* Location/Qualifiers

1. 265074

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-19984"

1. 2674

/note="wgs end extension

clone end: T7"

complement(5474..5908)

/note="clone boundary

site:BCORI

clone\_end: T7

end sequence: BH354676"

complement(255145..258153)

/note="clone boundary

clone\_end: Sp6

site:BCORI

end sequence: BH354677"

257056..258127

/note="wgs end extension

clone\_end: Sp6"

258623..262317

/note="wgs end extension

clone\_end: Sp6"

ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 265074;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 CGACTGTGATCTCTCA 17

250504 CGACTGTGATCTCTCA 250520

RESULT 18

AC112826/c

LOCUS AC112826 267779 bp DNA linear HTG 08-OCT-2002

# DEFINITION

Rattus norvegicus clone CH230-21P15, \*\*\* SEQUENCING IN PROGRESS

\*\*\*

AC112826

AC112826.3 GI:22856778

VERSION

KEYWORDS

HTG; HTGS PHASE2; HTGS DRAFT; HTGS\_ENRICHED.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 267779)

Muzny,D.Marie, Metker,M.Lee, Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Albrooks,S., Amth,A., Anguiano,D.,

Anyalebechi,V., Ayoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bennet,F.,

Biewald,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., d'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Georgiev,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guerra,W.,

Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,

Hollins,B., Howells,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyet,A.,

Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorenshewa,L., Louissege,H., Lozano,R.J., Lu,X., Ma,J.,

Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,

Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,

Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Natr,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwokelemeh,O., Okumu,G., Olarunsaogun,A., Pal,S., Parks,K.,

Paeternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,

Plopper,F., Poindecker,A., Popovic,D., Primbs,B., Pu,L.,

Puazo,M., Quiroz,J., Rachlin,B., Reeves,K., Reiter,M.A., Reigh,R.,

Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,

Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,

Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,

Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Snajls,D.,

Sneed,A., Sodergren,B., Song,X.-Z., Sorrell,R., Sosa,J.,

Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejo,J., Umani,K.,

Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,

Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

Williams,G., Willson,R., Wiczyski,R., Wooden,H., Worley,K.,

Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von

Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,

Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 267779)

Worley,K.C.

Direct Submission

Submitted (25-FEB-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 267779)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (08-OCT-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

## COMMENT

Baylor Plaza, Houston, TX 77030, USA  
 On Sep 14, 2002 this sequence version replaced gi:21738648.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

## ----- Genome Center

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GRW  
 Center clone name: CH230-21P15

## ----- Summary Statistics

Assembly program: Phrap; version 0.990329  
 Consensus quality: 223471 bases at least Q40  
 Consensus quality: 225356 bases at least Q30  
 Consensus quality: 226704 bases at least Q20  
 Estimated insert size: 241663; sum-of-contigs estimation  
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 267779: contig of 267779 bp in length.  
 Location/Qualifiers

FEATURES  
 source  
 1..267779  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-21P15"  
 81136..83530  
 /note="wgs\_contig"

ORIGIN  
 misc\_feature

Query Match 94.4%; Score 17; DB 2; Length 267779;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCTCCAT 18  
 |||||  
 Db 90770 GACTGTGAATCTCCAT 90754

RESULT 19  
 AL627428 275838 bp DNA linear HTG 23-OCT-2002  
 LOCUS Mus musculus chromosome 11 clone RP23-386C14, 15 unordered pieces.  
 ACCESSION AL627428.23 GI:24366466  
 VERSION AL627428.23 HTG; HTGS PHASE1; HTGS CANCELLED.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 275838)

AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

Clark, S.  
 Direct Submission  
 Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
[humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk) Clone requests: [clonerequests@sanger.ac.uk](mailto:clonerequests@sanger.ac.uk)  
 On Oct 24, 2002 this sequence version replaced gi:20502228.

## ----- Genome Center

Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 ----- Project Information  
 Center project name: bmk82014

## ----- Summary Statistics

Assembly program: XGAP; version 4.5  
 Sequencing vector: plasmid; 108752; 100% of reads  
 Chemistry: Dye-terminator; 1% of reads  
 Chemistry: Dye-terminator Big Dye; 98% of reads  
 Consensus quality: 270886 bases at least Q40  
 Consensus quality: 272427 bases at least Q30  
 Consensus quality: 273542 bases at least Q20  
 Insert size: 274438; sum-of-contigs  
 Insert size: 211998; 4.9% error; agarose-fp  
 Quality coverage: 10.60x in Q20 bases; sum-of-contigs quality  
 coverage: 13.87x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 15 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 2049: contig of 2049 bp in length  
 2050 2149: gap of 100 bp  
 2150 7109: contig of 4960 bp in length  
 7110 7209: gap of 100 bp  
 7210 15565: contig of 8356 bp in length  
 15566 15665: gap of 100 bp  
 15666 27153: contig of 11488 bp in length  
 27154 27253: gap of 100 bp  
 27254 30924: contig of 3671 bp in length  
 30925 31024: gap of 100 bp  
 31025 37642: contig of 6618 bp in length  
 37643 37742: gap of 100 bp  
 37743 105267: contig of 67525 bp in length  
 105268 105367: gap of 100 bp  
 105368 251950: contig of 146583 bp in length  
 251951 252050: gap of 100 bp  
 252051 259747: contig of 7697 bp in length  
 259748 259847: gap of 100 bp  
 259848 262155: contig of 2308 bp in length  
 262156 262255: gap of 100 bp  
 262256 264502: contig of 2247 bp in length  
 264503 264602: gap of 100 bp  
 264603 266975: contig of 2373 bp in length  
 266976 267075: gap of 100 bp  
 267076 269446: contig of 2371 bp in length  
 269447 271564: gap of 100 bp  
 269547 271564: contig of 2018 bp in length  
 271565 271664: gap of 100 bp  
 271665 275838: contig of 4174 bp in length.

FEATURES  
 source

1..275838  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone="RP23-386C14"  
 /clone\_lib="RP23-386C14"  
 1..2049  
 /note="assembly\_fragment:07407"

```

misc_feature      fragment chain:1"
2150..7109
/note="assembly_fragment:07527
fragment_chain:1"
misc_feature      7210..15565
/note="assembly_fragment:07041
fragment_chain:1"
misc_feature      15666..27153
/note="assembly_fragment:07271
fragment_chain:1"
misc_feature      27254..30924
/note="assembly_fragment:07135
fragment_chain:2"
misc_feature      31025..37642
/note="assembly_fragment:06991
fragment_chain:2"
misc_feature      37743..105267
/note="assembly_fragment:03775"
fragment_chain:2"
misc_feature      105368..251950
/note="assembly_fragment:05020"
misc_feature      252051..259747
/note="assembly_fragment:05467"
misc_feature      259848..262155
/note="assembly_fragment:05550"
misc_feature      262256..264502
/note="assembly_fragment:05948"
misc_feature      264603..266975
/note="assembly_fragment:06013"
misc_feature      267076..269446
/note="assembly_fragment:06040"
misc_feature      269547..271564
/note="assembly_fragment:07428"
misc_feature      271665..275838
/note="assembly_fragment:07590"

```

## ORIGIN

```

Query Match      94.4% Score 17; DB 2; Length 275838;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 GACTGTGATCCTCCAT 18
Db      155872 GACTGTGATCCTCCAT 155888

```

```

RESULT 20
AC118955/AC118955/
LOCUS      AC118955
DEFINITION Rattus norvegicus clone CH230-21006, *** SEQUENCING IN PROGRESS
*** 4 unordered pieces.
AC118955      328868 bp      DNA      linear      HTG 09-OCT-2002
AC118955.6 GI:23269762
VERSION      HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
KEYWORDS     Rattus norvegicus (Norway rat)
SOURCE       Rattus norvegicus
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 328868)
Munzy D.Marie, Metzger M.Lee, Abramson S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguliano D.,
Anyalebech V., Aoyagi A., Ayodeji M., Baca B., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biewald K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Buttrill K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Coyle M., Cree A., D'Souza L.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denison S., Deramo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Esecotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,

```

```

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL
TITLE
AUTHORS
REFERENCE
JOURNAL

```

## COMMENT

Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M., Gebregregis B., Geer K., Gill R., Grady M., Guerra W., Guevara W., Gunaratne P., Haaland W., Hamill C., Hamilton C., Hamilton K., Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J., Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M., Hollins B., Howells S., Hulik S., Hume J., Idlibit D., Jackson A., Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolyet A., Karpach S., Kelly S., Khan Z., King L., Kovar C., Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J., Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J., Lorensuhewa L., Louisedge H., Lozada R.J., Lu X., Ma J., Maheshwari M., Mahindartne M., Mahmoud M., Malloy K., Mangum A., Mangum B., Mapua P., Martin K., Martin R., Martinez B., Mashiney S., McLeod M.P., McNeill T.Z., Meenen B., Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S., Morgan M., Morris K., Morris S., Mundasa M., Murphy M., Natl L., Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Nockelemech O., Okwoum G., Olarnpungoon A., Pal S., Parks K., Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C., Plopper F., Poindexter A., Popovic D., Prims E., Pu L.L., Puzo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R., Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F., Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J., Sanders M., Savery G., Scherer S., Scott G., Shatman S., Shen H., Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajls D., Sneed A., Sodergren E., Song X.-Z., Sorelle R., Soosa J., Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C., Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Uমান K., Valas R., Vera V., Villaseana D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren J., Warren J., Wei X., White F., Williams G., Willson R., Wiczyski R., Wooden H., Worley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Zhou D., von Weinstock G., and Gibbs R.A.

Unpublished  
2 (bases 1 to 328868)

Direct Submission  
Submitted (22-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 328868)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (09-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 23, 2002 this sequence version replaced gi:21746759. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlantis/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: G02C  
Center clone name: CH230-21006  
Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 212214 bases at least Q40  
Consensus quality: 216733 bases at least Q30  
Consensus quality: 218850 bases at least Q20  
Estimated insert size: 256170; sum-of-contigs estimation  
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.bjsgc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.bjsgc.bcm.tmc.edu/docs/Genbank_draft_data.html))  
 \* NOTE: This sequence may represent more than one clone.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 5400: contig of 5400 bp in length  
 \* 5401 5500: gap of unknown length  
 \* 5501 296878: contig of 291278 bp in length  
 \* 296879 327503: gap of unknown length  
 \* 327504 327603: contig of 30625 bp in length  
 \* 327604 328868: contig of 1265 bp in length.

FEATURES  
 source  
 1..328868  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-21006"  
 1..1272  
 /note="wgs\_contig"  
 1323..2986  
 /note="wgs\_contig"  
 3856..5400  
 /note="wgs\_contig"

ORIGIN  
 Query Match 94.4%; Score 17; DB 2; Length 328868;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18  
 |||||  
 Db 4207 GACTGTGAATCCTCCAT 4191

RESULT 21  
 MMANT02/c 84 bp DNA linear ROD 10-FEB-1997  
 LOCUS  
 DEFINITION Murine gene fragment for cellular tumour antigen p53 (exon 2).  
 ACCESSION X00876  
 VERSION X00876.1 GI:871420  
 KEYWORDS antigen.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D.  
 A single gene and a pseudogene for the cellular tumour antigen p53  
 Nature 306 (5943), 594-597 (1983)  
 JOURNAL MEDLINE 84068204  
 PUBMED 6646235  
 2 (bases 1 to 83)  
 Bienz, B., Zakut-Houri, R., Givol, D. and Oren, M.  
 Analysis of the gene coding for the murine cellular tumour antigen  
 p53  
 EMBO J. 3 (9), 2179-2183 (1984)  
 JOURNAL MEDLINE 85027173  
 PUBMED 6092064  
 On Jun 23, 1995 this sequence version replaced gi:43974.  
 FEATURES  
 location/Qualifiers  
 1..84  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 join(X00875.1:1..536.1..84,X00877.1:1..21,X00878.1:1..261,  
 X00879.1:1..184,X00880.1:1..113,X00881.1:1..110,

## CDS

X00882.1:1..137,X00883.1:1..74,X00884.1:1..107,  
 X00885.1:1..522)  
 /product="cellular tumour antigen p53"  
 /label=p53 mRNA  
 join(1..84,X00877.1:1..21,X00878.1:1..261,X00879.1:1..184,  
 X00880.1:1..113,X00881.1:1..110,X00882.1:1..137,  
 X00883.1:1..74,X00884.1:1..107,X00885.1:1..82)  
 /codon\_start=1  
 /label=p53 CDS  
 /product="cellular tumour antigen p53"  
 /protein\_id="CA25420.1"  
 /db\_xref="GI:871421"  
 /db\_xref="GOA:P02340"  
 /db\_xref="Swiss-Prot:P02340"  
 /translation="MTAMERSQSDISLELPISQETFSGLWKLPPEDILPSHCMDDL  
 LIPDVBFEFGPEBALVSGAPAPADPVPAPAPATPPLSPFPSPKTYOG  
 NYGHLFLQSGAKSVMTYSPPLNTFCOLVYTCVQALWVSTPPAGSRVRAALY  
 KKSQMTVEVVRCPHRCGSDGDLAPPHLIRVGNLYPEYLEDROTFFHVVVPE  
 PPSGSEYTIHYVMCNSSCMGMMNRRLITITLTPDSSGNLGRSPFVRCACG  
 RDRTERENRKKKVLCPPLPPGAKRALPTCSASFPQKKPLDGYFTLTKRGR  
 FEMREINLEALDELDAHATESGDSRAHSSYLTKTKQSTSRHKTKVKKVGPDS"  
 1..84  
 /label=ex2

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 84;  
 Best Local Similarity 94.4%; Pred. No. 2.4e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
 |||||  
 Db 27 CGACTGTGAATCCTCCAT 10

RESULT 22  
 AX435258/c 390 bp DNA linear PAT 26-JUN-2002  
 LOCUS  
 DEFINITION Sequence 3673 from Patent WO0229113.  
 ACCESSION AX435258  
 VERSION AX435258.1 GI:21660066  
 KEYWORDS  
 SOURCE Bacillus licheniformis  
 ORGANISM Bacillus licheniformis  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 1  
 Berka, R. and Clausen, I.G.  
 Methods for monitoring multiple gene expression  
 Patent: WO 0229113-A 3673 11-Apr-2002;  
 Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)  
 FEATURES  
 location/Qualifiers  
 1..390  
 /organism="Bacillus licheniformis"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:11402"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 390;  
 Best Local Similarity 94.4%; Pred. No. 2.2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
 |||||  
 Db 120 CGACTGTGAATCCTCCAT 103

RESULT 23  
 AB038502/c 836 bp mRNA linear INV 19-AUG-2000  
 LOCUS  
 DEFINITION Caenorhabditis elegans mRNA for galelectin LEC-11, complete cds.  
 ACCESSION AB038502  
 VERSION AB038502.1 GI:9857640  
 KEYWORDS galelectin LEC-11; lec-11.  
 SOURCE Caenorhabditis elegans

ORGANISM *Caenorhabditis elegans*  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita;  
 Rhabditidae; Rhabditidae; Pelodermidae; *Caenorhabditis*.  
 REFERENCE 1 (sites)  
 AUTHORS Hirabayashi, Y., Hayama, K. and Kasai, K.  
 TITLE Novel galactins found in *C. elegans*  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 836)  
 AUTHORS Hirabayashi, Y., Hayama, K. and Kasai, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-FEB-2000) Jun Hirabayashi, Faculty of Pharmaceutical Sciences, Teikyo University, Department of Biological Chemistry, Suetsushi 1091-1, Sagami-cho, Kanagawa 199-0195, Japan (E-mail: j-hirabayashi@pharm.teikyo-u.ac.jp, Tel:81-426-85-3741, Fax:81-426-85-3742)  
 FEATURES  
 source Location/Qualifiers  
 1..836  
 /organism="Caenorhabditis elegans"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6239"  
 gene 1..836  
 /gene="lec-11"  
 mlec\_rna 1..22  
 /gene="lec-11"  
 /note="SL1 trans-spliced leader sequence"  
 CDS 97..795  
 /gene="lec-11"  
 /codon\_start=1  
 /product="galactin LEC-11"  
 /protein\_id="BAB11967.1"  
 /db\_xref="GI:9857641"  
 /translation="MAYIENPPIPLTPIYEGKREHSSVIEHGEVFGPGGPTVEP  
 PTNGVALHISVNGSYGQNTIVFNHLQKGMHREHQNITMRPFCRIHNEHK  
 YSHVDGHHIGHHKKPKRIAMATVAGDRLVGRIHFNPKHNGGTETIGVQMVP  
 GPTIVVQPPPMIVPQIVPVVQIMQMPVQIPSPVPIVPIATPIIYQPEV  
 IYMDGYHHRHHH"  
 822..827  
 /gene="lec-11"  
 /note="1ec-11"  
 ORIGIN  
 polyA\_signal  
 Query Match 91.1%; Score 16.4; DB 3; Length 836;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTAATCTCCAT 18  
 |||||  
 Db 223 CGACTGTAATCTCCCT 206  
 |||||  
 RESULT 24  
 MMU297973/c 1096 bp mRNA linear ROD 24-NOV-2000  
 LOCUS MMU297973  
 DEFINITION Mus musculus mRNA for transformation related protein 53 (Trp53 gene).  
 ACCESSION AJ297973  
 VERSION AJ297973.1 GI:11342598  
 KEYWORDS transformation related protein 53; Trp53 gene.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS Rothbarth, K. and Werner, D.  
 TITLE Proteasome-mediated degradation antagonizes critical levels of the  
 JOURNAL apoptosis-inducing CID protein  
 REFERENCE 2 (bases 1 to 1096)  
 AUTHORS Werner, D.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-NOV-2000) Werner D., Biochemistry of the Cell, German  
 JOURNAL Cancer Research Center, Im Neuenheimer Feld 280, Heidelberg, 69120,  
 GERMANY  
 FEATURES  
 Location/Qualifiers

source 1..1096  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="EK398"  
 /tissue type="Ehrlich ascites tumor"  
 gene 1..1096  
 /gene="p53"  
 CDS 1..924  
 /gene="p53"  
 /note="Premature termination due to the deletion (compared to wildtype)"  
 /product="transformation related protein 53"  
 /codon\_start=1  
 /protein\_id="CAC17147.1"  
 /db\_xref="GI:11342599"  
 /db\_xref="GOA:Q9ER40"  
 /db\_xref="TrEMBL:Q9ER40"  
 /translation="MSAMRESQSDISLEPLSOTFSGIWKILPPEDILSPHGMDDL  
 LLPDVERPFGESBALRVSGAPAAQDPVTETPGPVAPATPMPPLSPFSQXTYOG  
 NYGHLGPILOSGLKSVCTYSPLNKLFCOLATCPVQLVSAITPFGSRVRLMALT  
 KKSQMTVEVARCPHHRCSDDGLAPQHLIRVEGNLYPEHLEDROTFRHSVVVPR  
 PPEAGSEYTTIHYKMCNNSCMGMMNRPIITITLEDSSGNLIGRDSFEVRVACPG  
 RDRRTBEENFRKKEVLCPPLPGSAKDPRA"  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 10; Length 1096;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTAATCTCCAT 18  
 |||||  
 Db 27 CGACTGTAATCTCCAT 10  
 |||||  
 RESULT 25  
 AR052877/c 1173 bp DNA linear PAT 29-SEP-1999  
 LOCUS AR052877  
 DEFINITION Sequence 214 from patent US 5833975.  
 ACCESSION AR052877  
 VERSION AR052877.1 GI:5977739  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1173)  
 AUTHORS Paoletti, B., Tartaglia, J. and Cox, W. I.  
 TITLE Canarypox virus expressing cytokine and/or tumor-associated antigen  
 JOURNAL DNA sequence  
 REFERENCE Patent: US 5833975-A 214 10-NOV-1998;  
 LOCATION/Qualifiers  
 source 1..1173  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 6; Length 1173;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTAATCTCCAT 18  
 |||||  
 Db 27 CGACTGTAATCTCCAT 10  
 |||||  
 RESULT 26  
 AR127895/c 1173 bp DNA linear PAT 16-MAY-2001  
 LOCUS AR127895  
 DEFINITION Sequence 45 from patent US 6183752.  
 ACCESSION AR127895  
 VERSION AR127895.1 GI:14115557  
 KEYWORDS  
 SOURCE Unknown.

ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 1173)  
AUTHORS Epstein,S.B., Finkel,T., Speir,E., Zhou,Y.Fu., Zhu,J., Erdlie,L.  
and Pincus,S.  
TITLE Restenosis/atherosclerosis diagnosis, prophylaxis and therapy  
JOURNAL Patent: US 6183752-A 45 06-FEB-2001;  
FEATURES Location/Qualifiers  
source 1..1173  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 6; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 27 CGACTGTGACTCTCCAT 10  
|||||  
|||||

RESULT 27  
AR288189/c 1173 bp mRNA linear PAT 12-JUN-2003  
LOCUS AR288189  
DEFINITION Sequence 214 from patent US 6537594.  
ACCESSION AR288189  
VERSION AR288189.1 GI:31675468  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 1173)  
Paoletti,E., Tartaglia,J. and Cox,W.I.  
TITLE Vaccina virus comprising cytokine and/or tumor associated antigen  
JOURNAL Patent: US 6537594-A 214 25-MAR-2003;  
FEATURES Location/Qualifiers  
source 1..1173  
/organism="unknown"  
/mol\_type="mRNA"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 6; Length 1173;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 27 CGACTGTGACTCTCCAT 10  
|||||  
|||||

RESULT 28  
AY044188/c 1176 bp mRNA linear ROD 21-DEC-2001  
LOCUS AY044188  
DEFINITION Mus musculus transformation related protein 53 (Trp53) mRNA,  
complete cds.  
ACCESSION AY044188  
VERSION AY044188.1 GI:15375071  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1176)  
Irminger-Finger,I., Leung,W.-C., Li,J., Dubois-Dauphin,M., Harb,J.,  
Feki,A., Jefford,C.B., Soriano,J.V., Deconh,W., Montesano,R. and  
Krause,K.-H.  
TITLE Identification of BARD1 as Mediator between Proapoptotic Stress and  
p53-Dependent Apoptosis  
JOURNAL Mol. Cell 8 (6), 1255-1266 (2001)  
PUBMED 11779501  
REFERENCE 2 (bases 1 to 1176)

AUTHORS Irminger-Finger,I., Feki,A., Jefford,C.B. and Molnarfi,N.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUL-2001) Geriatrics, Univ. of Geneva, 2 ch.  
Petit-Bel-Air, Geneva CH-1225, Switzerland  
FEATURES Location/Qualifiers  
source 1..1176  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/cell\_line="FAC-2"  
/cell\_type="epithelial"  
/tissue\_type="mammary gland"  
1..1176  
/gene="Trp53"  
1..1176  
/gene="Trp53"  
/note="p53; tumor suppressor"  
/codon\_start=1  
/product="transformation related protein 53"  
/protein\_id="P4K94783.1"  
/db\_xref="GI:15375072"  
/translation="MTAMBSQSDISLEPLSQETFSGLWKLPPEDILPSPHCMDDL  
LPDQVEFPFGSEALRVSGAPADPVTETGPVAPAPATPPISSFPQKTOG  
NYGPHLGIQSGTAKSYMCTVSPINKLFCOLAKTQVQIMVSAATPAGSRVAMATY  
KKSQHWREYVRCPHHRCSDDGGLAPPHLITVEGNLYREYLEDQTPHRSVIVPYE  
PPFASRYTHIKYMCNSCWGGMRRPILITITLBDSSGNLGRDSFEVRVACAPG  
RDMRTBERENRKEVLCEPLPGSARALPTCTASAPPOKKRPLDGEYFTLKYGRRR  
FEMREILNMLRLKDAHATBESGDSRAHSYLTKKKGQSTSRHKTMVKKVGPDDSD"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 10; Length 1176;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
Db 27 CGACTGTGACTCTCCAT 10  
|||||  
|||||

RESULT 29  
AF161020/c 1183 bp mRNA linear ROD 18-JUL-1999  
LOCUS AF161020  
DEFINITION Mus musculus tumor suppressor p53 (p53) mRNA, complete cds.  
ACCESSION AF161020  
VERSION AF161020.1 GI:5524685  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1183)  
Jiang,G.C., Yuan,L.Z., Wei,K. and Guo,X.M.  
TITLE Sequence of tumor suppressor gene p53 in radiosensitive mouse breast  
carcinoma cell line Sx-9  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1183)  
AUTHORS Jiang,G.C., Yuan,L.Z., Wei,K. and Guo,X.M.  
TITLE Direct Submission  
JOURNAL Submitted (21-JUN-1999) Institute of Radiation Medicine, Academy of  
Military Medical Sciences, No.27 Taiping Road, Haiding District,  
Beijing 100850, P. R. China  
FEATURES Location/Qualifiers  
source 1..1183  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/cell\_line="Sx-9"  
/tissue\_type="radiosensitive breast carcinoma"  
1..1183  
/gene="p53"  
11..1183  
/gene="p53"  
/codon\_start=1



/product="tumor suppressor p53"  
 /protein\_id="A044340.1"  
 /db\_xref="GI:5524686"  
 /translation="MTAMEBSQSDISLEPLSQETPSGLWKLPPEDILPSFHCMDLL  
 LLPDVEBFEGPSBALVSGAPAAQDPVETPGVAPAPATPWLSPFVSQKTYQG  
 NYGFLGFLQSTAKSVCTSPPLNKLFCQALATCPVQLMVSAATPPAGSRVAMAIY  
 KKSQMTVEVRACPHHRCSDGDLAPROH.IRVGNLYPEYLEDROTFRRSVVAYE  
 PPKASRTTTHYKTMCSNCSGGMRRPILITITLBSNGLGRDSFEVRVACRG  
 RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLTKGRK  
 FEMRELNEALELKDAAHTBSGDSRAHSYLTKKQGSTSRHKTKVKKVGPDS"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1103;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGTGATCTCCAT 18  
 |||||  
 Db 37 CGACTGTGATCTCCAT 20

RESULT 30  
 AF051368 1208 bp mRNA linear ROD 16-MAR-1998  
 LOCUS AF051368  
 DEFINITION Mus musculus tumor suppressor p53 mRNA, complete cds.  
 ACCESSION AF051368  
 VERSION AF051368.1 GI:2961246  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1208)  
 Frostejo, L., Nilsson, J., Wandzioch, E. and Heby, O.  
 Sequence of tumor suppressor p53 in mouse lymphoid leukemia cell  
 line L1210

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 2 (bases 1 to 1208)  
 Frostejo, L., Nilsson, J., Wandzioch, E. and Heby, O.  
 Direct Submission  
 Submitted (27-FEB-1998) Dept. Cellular and Developmental Biology,  
 Umea University, Hus 1, Umea 901 87, Sweden  
 Location/Qualifiers

## FEATURES

source

1..1208  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /cell\_line="L1210"  
 /tissue\_type="lymphoid leukemia"  
 26..1198  
 /note="MLP53; mutant"  
 /codon\_start=1  
 /product="tumor suppressor p53"  
 /protein\_id="AAC05704.1"  
 /db\_xref="GI:2961247"  
 /translation="MTAMEBSQSDISLEPLSQETPSGLWKLPPEDILPSFHCMDLL  
 LLPDVEBFEGPSBALVSGAPAAQDPVETPGVAPAPATPWLSPFVSQKTYQG  
 NYGFLGFLQSTAKSVCTSPPLNKLFCQALATCPVQLMVSAATPPAGSRVAMAIY  
 KKSQMTVEVRACPHHRCSDGDLAPROH.IRVGNLYPEYLEDROTFRRSVVAYE  
 PPKASRTTTHYKTMCSNCSGGMRRPILITITLBSNGLGRDSFEVRVACRG  
 RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLTKGRK  
 FEMRELNEALELKDAAHTBSGDSRAHSYLTKKQGSTSRHKTKVKKVGPDS"  
 263..266  
 /note="encodes substitution of alanine for wildtype  
 valine"

## CDS

misc\_feature  
 /note="encodes substitution of alanine for wildtype  
 valine"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1208;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGTGATCTCCAT 18  
 |||||

Db 52 CGACTGTGATCTCCAT 35

RESULT 31  
 M05P53B/c 1241 bp mRNA linear ROD 27-APR-1993  
 LOCUS M05P53B  
 DEFINITION Mouse p53 mRNA, complete cds, clone p53-m11.  
 ACCESSION M13873  
 VERSION M13873.1 GI:200200  
 KEYWORDS  
 SOURCE p53 gene.  
 ORGANISM Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1241)  
 Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and  
 Rotter, V.  
 Immunologically distinct p53 molecules generated by alternative  
 splicing

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Mol. Cell. Biol. 6 (9), 3232-3239 (1986)  
 MEDLINE  
 PUBMED  
 3023970  
 COMMENT Original source text: Mouse (BALB/c) Meth A library, cDNA to mRNA,  
 clone p53-m11.  
 FEATURES  
 source

## FEATURES

source

Location/Qualifiers  
 1..1241  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /sub\_species="domesticus"  
 /db\_xref="taxon:10090"  
 /map="11"  
 /tissue\_lib="Meth A"  
 1..1241  
 /gene="p53"  
 69..1241  
 /gene="p53"  
 /codon\_start=1  
 /protein\_id="AAA39882.1"  
 /db\_xref="GI:200201"  
 /translation="MTAMEBSQSDISLEPLSQETPSGLWKLPPEDILPSFHCMDLL  
 LLPDVEBFEGPSBALVSGAPAAQDPVETPGVAPAPATPWLSPFVSQKTYQG  
 NYGFLGFLQSTAKSVCTSPPLNKLFCQALATCPVQLMVSAATPPAGSRVAMAIY  
 KKSQMTVEVRACPHHRCSDGDLAPROH.IRVGNLYPEYLEDROTFRRSVVAYE  
 PPKASRTTTHYKTMCSNCSGGMRRPILITITLBSNGLGRDSFEVRVACRG  
 RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLTKGRK  
 FEMRELNEALELKDAAHTBSGDSRAHSYLTKKQGSTSRHKTKVKKVGPDS"

## gene

CDS

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1241;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGTGATCTCCAT 18  
 |||||  
 Db 95 CGACTGTGATCTCCAT 78

RESULT 32  
 M05P53A/c 1285 bp mRNA linear ROD 27-APR-1993  
 LOCUS M05P53A/c  
 DEFINITION Mouse p53 mRNA, complete cds, clone pc053.  
 ACCESSION M13872  
 VERSION M13872.1 GI:200198  
 KEYWORDS  
 SOURCE p53 gene.  
 ORGANISM Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1285)  
 Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and  
 Rotter, V.  
 Immunologically distinct p53 molecules generated by alternative

REFERENCE  
 AUTHORS  
 TITLE  
 Immunologically distinct p53 molecules generated by alternative

JOURNAL Mcl. Cell. Biol. 6 (9), 3232-3239 (1986)  
 JOURNAL 87064640  
 PUBMED 3023970

COMMENT Original source text: Mouse (BALB/c) nontransformed helper T-cell,  
 CDNA to mRNA, clone pCD-p53.  
 FEATURES Location/Qualifiers

## SOURCE

## CDS

1..1285  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /sub\_species="domesticus"  
 /db\_xref="taxon:10090"  
 /cell\_type="nontransformed helper T-cell"  
 113..1285  
 /codon\_start=1  
 /product="p53 protein"  
 /protein\_id="AAA39881.1"  
 /db\_xref="GI:200199"  
 /translation="MTAMESOSDISLEPLISOFTSGMLKLLPREDILSPHCMDL  
 LLPODVEFFEGPSEALRVSGAPAAQDPVTETPGVAPAPATPWPSSVPSQKTYOG  
 NYGHLGFLQSGTAKSVCTYSPPPLNKLFCQIAKTCVQLVWSATPPGSRVAMAIY  
 KKSQHTVEVRRCPHHERCSDGDLAPPHLIRVEGNLYPEYLEDROTFRHSVVVPE  
 PPEAGSEYTIHYKVMCNSSCMGGMNRPIITITLEDSSGNILGRDSFEVAVCAFG  
 RDRRTSEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKKPLDGEYFTLKIRGR  
 FEMFREINALELKDANATEBSGDSRAHSYLTKTKGQTSIRHKKTWKKVGPDS"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1285;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18  
 |||||  
 Db 139 CGACTGGAATCCTCCAT 122

RESULT 33  
 AB017816/c 1305 bp mRNA linear ROD 09-JUL-1999  
 LOCUS Mus musculus mRNA for p53, complete cds.  
 AB017816  
 VERSION AB017816.1 GI:5421812  
 KEYWORDS p53.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (sites)  
 AUTHORS Araki, R., Fukumura, R., Fujimori, A., Tatsumi, K. and Abe, M.  
 TITLE Characterization of DNA-PKcs null mutant sx9  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1305)  
 Fujimori, A. and Abe, M.  
 DIRECT SUBMISSION  
 Submitted (22-SEP-1998) Masumi Abe, National Institute of  
 Radiological Sciences, Dept. of Biology and Oncology, Anagawa  
 4-9-1, Inage-ku, Chiba, Chiba 263-8555, Japan  
 (E-mail: abemasumi@nirs.go.jp, Tel:043-206-3219,  
 Fax:043-251-4593)

## FEATURES

## SOURCE

1..1305  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /cell\_line="sx9"  
 /cell\_type="epithelial"  
 /tissue\_type="mammary carcinoma"  
 102..1274  
 /codon\_start=1  
 /product="p53"  
 /protein\_id="BAA82340.1"  
 /db\_xref="GI:5421813"

## CDS

/translation="MTAMESOSDISLEPLISOFTSGMLKLLPREDILSPHCMDL  
 LLPODVEFFEGPSEALRVSGAPAAQDPVTETPGVAPAPATPWPSSVPSQKTYOG  
 NYGHLGFLQSGTAKSVCTYSPPPLNKLFCQIAKTCVQLVWSATPPGSRVAMAIY  
 KKSQHTVEVRRCPHHERCSDGDLAPPHLIRVEGNLYPEYLEDROTFRHSVVVPE  
 PPEAGSEYTIHYKVMCNSSCMGGMNRPIITITLEDSSGNILGRDSFEVAVCAFG  
 RDRRTSEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKKPLDGEYFTLKIRGR  
 FEMFREINALELKDANATEBSGDSRAHSYLTKTKGQTSIRHKKTWKKVGPDS"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1305;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18  
 |||||  
 Db 128 CGACTGGAATCCTCCAT 111

RESULT 34  
 M13874/c 1322 bp mRNA linear ROD 27-APR-1993  
 LOCUS Mouse p53 mRNA, complete cds, clone p53-m8.  
 DEFINITION M13874  
 ACCESSION M13874.1 GI:200202  
 VERSION M13874.1 GI:200202  
 KEYWORDS p53 gene.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1322)  
 AUTHORS Araki, N., Nomura, D., Yokota, K., Wolf, D., Bill, E., Shohat, O. and  
 Rotter, V.  
 TITLE Immunologically distinct p53 molecules generated by alternative  
 splicing  
 JOURNAL Mcl. Cell. Biol. 6 (9), 3232-3239 (1986)  
 MEDLINE 87064640  
 PUBMED 3023970  
 REFERENCE 2 (sites)  
 AUTHORS Han, K.A. and Kulesz-Martin, M.F.  
 TITLE Alternatively spliced p53 RNA in transformed and normal cells of  
 different tissue types  
 JOURNAL Nucleic Acids Res. 20 (8), 1979-1981 (1992)  
 MEDLINE 92253421  
 PUBMED 1579500

COMMENT Original source text: Mouse Ab-MuLV transformed Meth A fibroblast  
 and normal T-cell line L12, CDNA to mRNA, clone p53-m8.  
 FEATURES Location/Qualifiers

## SOURCE

## gene

## CDS

1..1322  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /map="11"  
 /cell\_line="normal T-cell line L12"  
 /cell\_type="Ab-MuLV transformed Meth A fibroblast"  
 1..1322  
 /gene="p53"  
 /gene="p53"  
 /codon\_start=1  
 /protein\_id="AAA39883.1"  
 /db\_xref="GI:200203"  
 /translation="MTAMESOSDISLEPLISOFTSGMLKLLPREDILSPHCMDL  
 LLPODVEFFEGPSEALRVSGAPAAQDPVTETPGVAPAPATPWPSSVPSQKTYOG  
 NYGHLGFLQSGTAKSVCTYSPPPLNKLFCQIAKTCVQLVWSATPPGSRVAMAIY  
 KKSQHTVEVRRCPHHERCSDGDLAPPHLIRVEGNLYPEYLEDROTFRHSVVVPE  
 PPEAGSEYTIHYKVMCNSSCMGGMNRPIITITLEDSSGNILGRDSFEVAVCAFG  
 RDRRTSEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKKPLDGEYFTLKIRGR  
 FEMFREINALELKDANATEBSGDSRAHSYLTKTKGQTSIRHKKTWKKVGPDS"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1322;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**Qy** 1 CGACTGGAATCTCCAT 18  
 |||||  
**Db** 80 CGACTGGAATCTCCAT 63  
 |||||

**RESULT 35**  
**LOCUS** AB017815 1360 bp mRNA linear ROD 09-JUL-1999  
**DEFINITION** Mus musculus mRNA for p53, complete cds.  
**ACCESSION** AB017815  
**VERSION** AB017815.1 GI:5421807  
**KEYWORDS** p53.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
**REFERENCE** Araki, R., Fukumura, R., Fujimori, A., Tatematsu, K. and Abe, M.  
**TITLE** Characterization of DNA-PKcs null mutant SK9  
**JOURNAL** Unpublished  
**AUTHORS** Fujimori, A. and Abe, M.  
**REFERENCE** Submitted (22-SEP-1998) Masumi Abe, National Institute of  
**JOURNAL** Radiological Sciences, Dept. of Biology and Oncology, Anagawa  
 4-9-1, Inage-Ku, Chiba, Chiba 263-8555, Japan  
 (E-mail: abemasum@nres72.nirs.go.jp, Tel:043-206-3219,  
 Fax:043-251-4593)  
**FEATURES** Location/Qualifiers  
 source 1..1360  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /cell\_line="SR-1"  
 /cell\_type="epithelial"  
 /tissue\_type="mammary carcinoma"  
 102..1374  
 /codon\_start=1  
 /product="p53"  
 /protein\_id="BA02339.1"  
 /db\_xref="GI:5421808"  
 /translation="MTMERSQSDISLEPLSOFTFSGLMKLLPPEDLPSPHCMDDL  
 LIPDVEPPEPGEALRVSGAPADPVTETPGWAPAPATPWPISFVPSQKTYOG  
 NYGHLGFLSGTAKSVCTTSPPLKLFQCLATCPVQLVSAATPPGSRVRAATIK  
 KSGQMTVEVRCPEHRCSDGDIAPQHLRVGNLYPEYLRQTFRHSVVVPEP  
 PEAGSEYTIHYKTMCNSSCMGMNRPIITITLLEDSGNLGRDSFEVAVACRG  
 DRRTBEENFRKEVLCELPFGSAKRALPTCTASPPQKKPLDGEYFTLKIRGRK  
 FEMRELNEALEMDAHTESGDSRAHSYLTKKKGQSTSRHKKTWVKVGPDS"  
**ORIGIN**  
 Query Match 91.1%; Score 16.4; DB 10; Length 1360;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**Qy** 1 CGACTGGAATCTCCAT 18  
 |||||  
**Db** 128 CGACTGGAATCTCCAT 111  
 |||||

**RESULT 36**  
**LOCUS** MPM53 1377 bp mRNA linear ROD 12-SEP-1993  
**DEFINITION** Mouse mRNA for transformation associated protein p53.  
**ACCESSION** X00741  
**VERSION** X00741.1 GI:53570  
**KEYWORDS** oncogene.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**TITLE** 1 (bases 1 to 1377)  
**AUTHORS** Jenkins, J.R., Rudge, K., Redmond, S. and Wade-Evans, A.

**TITLE** Cloning and expression analysis of full length mouse cDNA sequences  
 encoding the transformation associated protein p53  
**JOURNAL** Nucleic Acids Res. 12 (14), 5609-5626 (1984)  
**MEDLINE** 84272240  
**PUBMED** 6379601  
**COMMENT** Data kindly reviewed (19-FEB-1986) by A. Wade-Evans.  
**FEATURES** Location/Qualifiers  
 source 1..1377  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 123..1292  
 /note="unnamed protein product; p53"  
 /codon\_start=1  
 /protein\_id="CAA25323.1"  
 /db\_xref="GI:53571"  
 /db\_xref="GOA:P02340"  
 /db\_xref="Swiss-Prot:P02340"  
 /translation="MTMERSQSDISLEPLSOFTFSGLMKLLPPEDLPSPHCMDDL  
 LIPDVEPPEPGEALRVSGAPADPVTETPGWAPAPATPWPISFVPSQKTYOG  
 NYGHLGFLSGTAKSVCTTSPPLKLFQCLATCPVQLVSAATPPGSRVRAATIK  
 KSGQMTVEVRCPEHRCSDGDIAPQHLRVGNLYPEYLRQTFRHSVVVPEP  
 PEAGSEYTIHYKTMCNSSCMGMNRPIITITLLEDSGNLGRDSFEVAVACRG  
 DRRTBEENFRKEVLCELPFGSAKRALPTCTASPPQKKPLDGEYFTLKIRGRK  
 FEMRELNEALEMDAHTESGDSRAHSYLTKKKGQSTSRHKKTWVKVGPDS"  
**ORIGIN**  
 Query Match 91.1%; Score 16.4; DB 10; Length 1377;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**Qy** 1 CGACTGGAATCTCCAT 18  
 |||||  
**Db** 149 CGACTGGAATCTCCAT 132  
 |||||

**RESULT 37**  
**LOCUS** AF151353 1409 bp mRNA linear ROD 14-JUL-1999  
**DEFINITION** Mus musculus tumor suppressor p53 (p53) mRNA, complete cds.  
**ACCESSION** AF151353  
**VERSION** AF151353.1 GI:5081782  
**KEYWORDS** Mus musculus (house mouse)  
**SOURCE** Mus musculus  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE** Jimenez, G.S., Brynneson, F., Torres-Arzuaga, M.I., Priestley, A.,  
 Beeche, M., Saito, S., Sakaguchi, K., Appella, E., Jeggo, P.A.,  
 Taccioli, G.B., Wahl, G.M. and Hubank, M.  
**TITLE** DNA-dependent protein kinase is not required for the p53-dependent  
 response to DNA damage  
**JOURNAL** Nature 400 (6739), 81-83 (1999)  
**MEDLINE** 99330047  
**PUBMED** 10403253  
**REFERENCE** 2 (bases 1 to 1409)  
 Jimenez, G.S., Brynneson, F., Torres-Arzuaga, M.I., Priestley, A.,  
 Beeche, M., Saito, S., Sakaguchi, K., Appella, E., Jeggo, P.A.,  
 Taccioli, G.B., Wahl, G.M. and Hubank, M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (13-MAY-1999) Trafford Centre, University of Sussex,  
 Falmer, Brighton, Sussex BN1 9RY, United Kingdom  
**FEATURES** Location/Qualifiers  
 source 1..1409  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /map="39.0 cM"  
 /cell\_line="SGR11"  
 /cell\_type="fibroblast"  
 1..1409  
 gene

CDS

/gene="p53"  
123..1295  
/gene="p53"  
/codon\_start=1  
/product="tumor suppressor p53"  
/protein\_id="AAD39535.1"  
/db\_xref="GI:5081783"

variation

/translation="MTAMBSQSDISLEPLSGQETSGIMKLLPPEDILPSPHCMDDL  
LLPDVEFFEGSEALRVSGAPAAQDPVTEITPGVAPAPATPMPPLSFVPSQKYQG  
NYGPHLGFLOSSTAKSVCTYSPPLNKLFCOLAKTCVOLWVSATPPAGSRVAMAIY  
KKSQMTVEVRRCPEHRCSDGDLAPQHLIRVGNLYPEYLEDROTFRRSVVAVPE  
PEAGSEYTTIHYKVMCNSSCMGGMNRRLITLTLEDDSGNLLGRDSFEVRVACPG  
RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASAPPOKKKPLDGEYFTLKIRGRK  
FEMFRELNEALELKDAAHATESGDSRAHSSYLTKTKGQSTSRHKTKWKKVGPDS"  
694  
/gene="p53"  
/replace="t"

ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1409;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 149 CGACTGTGACTCTCCAT 132  
|||||

RESULT 38

AB021961/c 1429 bp mRNA linear ROD 14-APR-2000

LOCUS AB021961  
DEFINITION Mus musculus mutant p53 mRNA, complete cds.  
ACCESSION AB021961  
VERSION AB021961.1 GI:5421849  
KEYWORDS P53.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1429)  
Araki, R., Fukumura, R., Fujimori, A., Tatesumi, K. and Abe, M.  
Cell cycle in DNA-PKcs knock-out mice  
Unpublished  
2 (bases 1 to 1429)  
Fujimori, A. and Abe, M.  
Direct Submission  
Submitted (28-DEC-1998) Masumi Abe, National Institute of  
Radiological Sciences, Dept. of Biology and Oncology; Anagawa  
4-9-1, Inage-ku, Chiba, Chiba 263-8555, Japan  
(E-mail: abemasumi@nirs.go.jp, Tel:043-206-3219,  
Fax:043-251-4593)

FEATURES

source

1..1429  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="SCID"  
/db\_xref="taxon:10090"  
/cell\_line="SCGR-11"  
/cell\_type="fibroblast"  
/note="SCGR-11 cell is a derivative of SC33 fibroblast  
cell line derived from scid (severe combined immune  
deficiency) mice."  
1..1429  
/gene="p53"  
101..1273  
/gene="p53"  
/note="Amino acid no.191 leusine (L) in wild-type mouse  
p53 is substituted to arginine (R)."  
/codon\_start=1  
/product="p53"  
/protein\_id="BAA82344.1"  
/db\_xref="GI:5421850"  
/translation="MTAMBSQSDISLEPLSGQETSGIMKLLPPEDILPSPHCMDDL

variation

/gene="p53"  
/note="t in wild-type mouse p53"  
/replace="t"

ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1429;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18  
|||||  
Db 127 CGACTGTGACTCTCCAT 110  
|||||

RESULT 39

AB020317/c 1457 bp mRNA linear ROD 09-JUL-1999

LOCUS AB020317  
DEFINITION Mus musculus mRNA for p53, complete cds.  
ACCESSION AB020317  
VERSION AB020317.1 GI:5421827  
KEYWORDS P53.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1457)  
Fujimori, A. and Abe, M.  
P53 in 129-SVJ mice  
Published Only in Database (1999)  
2 (bases 1 to 1457)  
Fujimori, A. and Abe, M.  
Direct Submission  
Submitted (19-NOV-1998) Masumi Abe, National Institute of  
Radiological Sciences, Dept. of Biology and Oncology; Anagawa  
4-9-1, Inage-ku, Chiba, Chiba 263-8555, Japan  
(E-mail: abemasumi@nirs.go.jp, Tel:043-206-3219,  
Fax:043-251-4593)

FEATURES

source

1..1457  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129SVJ"  
/db\_xref="taxon:10090"  
/tissue\_type="lung fibroblast"  
1..1457  
/gene="p53"  
102..1274  
/gene="p53"  
/codon\_start=1  
/product="p53"  
/protein\_id="BAA82343.1"  
/db\_xref="GI:5421828"

ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1457;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18



/note="G is U in [1]"  
 /citation=[1]  
 1445. .1446  
 /note="UA is CU in [1]"  
 /citation=[1]  
 1447. .1448  
 /note="CU is UC in [1]"  
 /citation=[1]  
 1503  
 /note="U is A in [1]"  
 /citation=[1]  
 1558. .1559  
 /note="CU is UC in [1]"  
 /citation=[1]

Query Match 91.1%; Score 16.4; DB 10; Length 1773;  
 Best Local Similarity 94.4%; Pred. No. 2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18  
 184 CGACTGTGAATCTCCAT 167

RESULT 42 BC005448/c 1782 bp mRNA linear ROD 29-JUN-2004  
 LOCUS BC005448  
 DEFINITION Mus musculus transfection related protein 53, mRNA (CDNA clone  
 BC005448  
 VERSION BC005448.1 GI:13529427  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1782)  
 AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stadelman M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
 Schett P.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,  
 Carrinci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,  
 Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J.,  
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,  
 Wortley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Hellon E., Kettelman M., Madan A., Rodriguez S.,  
 Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,  
 Bouffard G.G., Rodriguez A.C., Grimwood J., Schmutz V., Myers R.M.,  
 Butlerfield J.S., Krzywinski M.I., Skalska U., Smalins D.E.,  
 Scherch A., Schein J.E., Jones S.J., and Marra M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE 1247932  
 JOURNAL Strausberg R.  
 PUBMED Direct Submission  
 REFERENCE Submitted (27-MAR-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
 USA

REMARK COMMENT  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-rc@mail.nih.gov](mailto:cgabs-rc@mail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcdeparil@stanford.edu](mailto:mcdeparil@stanford.edu)  
 Dickson, M., Schmutz V., Grimwood J., Rodriguez A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>  
 Series: IRAX Plate: 11 Row: 1 Column: 4  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 6755880.

FEATURES  
 source  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:6174 IMAGE:3599812"  
 /issue\_type="Mammary tumor. Metalloclonon-TGF alpha  
 model. 10 month old virgin mouse. Taken by biopsy."  
 /clone\_id="NCI CGAC\_Mam1"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1. .1782  
 /gene="Trp53"  
 /note="synonym: p53"  
 /db\_xref="locusid:22059"  
 /db\_xref="MGI:98834"  
 129. .1301  
 /gene="Trp53"  
 /codon\_start=1  
 /product="transformation related protein 53"  
 /protein\_id="AAH05448.1"  
 /db\_xref="GI:13529428"  
 /db\_xref="locusid:22059"  
 /db\_xref="MGI:98834"  
 /translation="MTMRSSQSDISLEPLPSQETPGSGWKLIPEDIIIPSHCHDDL  
 LHPQDVEFPFGSEALRVSGAPAAQDPTETPPGVPAPAPWPPLSPVQKYYOG  
 NYGFLFGFQSGAKSVCTSPPLKPCQAKTCPVQVWVSAAPPASRVANAIY  
 KKSQHTVEVVRRCPIHERCSDDGLAPPHLIRVEGNLYPEYLEDROTFRHSVVVYE  
 PPAGSGEYTVIRHKVNCSSCGMMRRPILITITLEDSSGNLGRDSFEVVCAPCG  
 RDRTEENPRKKEVCPPLPGSAKRALPTCTSAAPPKKKKPLDGEVETLKIRGRKR  
 FEFRRINALEIKDHAHIEBDSRAHSYTKTKKGQSTSHKKTWAKVGPDSO"

# ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1782;  
 Best Local Similarity 94.4%; Pred. No. 2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18  
 155 CGACTGTGAATCTCCAT 138

RESULT 43 K02110/c 2130 bp DNA linear ROD 10-MAY-2004  
 LOCUS K02110  
 DEFINITION Mus musculus tumor antigen p53 (p53) pseudogene, complete cds.  
 ACCESSION K02110 X01236  
 VERSION K02110.1 GI:200206  
 KEYWORDS antigen; p53 gene; processed pseudogene; tumor antigen.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2130)  
 AUTHORS Zakut-Houri R., Oren M., Bizen B., Lavie V., Hazun S. and Givol D.  
 TITLE A single gene and a pseudogene for the cellular tumour antigen p53  
 JOURNAL Nature 306 (5943), 594-597 (1983)  
 MEDLINE 84068204  
 PUBMED 6646235  
 REFERENCE 2 (bases 1 to 2130)

AUTHORS Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Glivol, D.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-APR-1993) Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot 76100, Israel  
 COMMENT On May 10, 2004 this sequence version replaced gi:53574.  
 FEATURES  
 source  
 1..2130  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /clone="pCH53-11"  
 <1..1826  
 /gene="p53"  
 /note="processed pseudogene"  
 /pseudo  
 1..185  
 /gene="p53"  
 /note="no similarity to p53 cDNA"  
 /pseudo  
 169..181  
 /rpt\_type=direct  
 262..1434  
 /gene="p53"  
 /note="tumor antigen p53"  
 /pseudo  
 /codon\_start=1  
 1807..1812  
 /gene="p53"  
 /note="putative"  
 1826  
 /gene="p53"  
 /note="putative"  
 1852..1864  
 /rpt\_family="direct"  
 repeat\_region  
 polyA\_signal  
 polyA\_site  
 1826  
 /gene="p53"  
 /note="putative"  
 1852..1864  
 /rpt\_family="direct"  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 10; Length 2130;  
 Best Local Similarity 94.4%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGACTGTGATCTCTCAT 18  
 |||||  
 Db 288 CGACTGTGATCTCTCAT 271  
 |||||  
 RESULT 44  
 AC023805/c  
 LOCUS AC023805  
 DEFINITION Mus musculus clone RP23-402K16, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 26  
 unordered pieces.  
 AC023805 86146 bp DNA linear HTG 16-OCT-2001  
 AC023805  
 VERSION AC023805.7 GI:16118068  
 KEYWORDS HTG; HTGS PHASE1.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 86146)  
 Metzer, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,  
 Decker, D., Thomas, S., Okumura, G., Carlock, C., Garner, T.,  
 Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,  
 Bunay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,  
 Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,  
 Fernandez, C., Ferraguto, D., Forcum-Taney, D., Gill, R.,  
 Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hognes, M.,  
 Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,  
 Kovari, C., Liu, J., Liu, W., Louisse, H., Lozano, R., Martin, R.,  
 Massey, B., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,  
 Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Oguri, M., Parlin, B.,  
 Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S.,  
 Williamson, A., Wrenford, G., Zhou, X., Bouck, J., Hodgson, A.,  
 Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein, G.,

Worley, K. and Gibbs, R.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 86146)  
 AUTHORS Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 COMMENT On Oct 14, 2001 this sequence version replaced gi:11079366.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: MABG  
 Center clone name: RP23-402K16  
 ----- Summary Statistics  
 Sequencing vector: M13; L08821  
 Chemistry: Dye-primer; Bodipy; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 80653 bases at least Q40  
 Consensus quality: 102703 bases at least Q40  
 Consensus quality: 112226 bases at least Q30  
 Estimated insert size: 102605; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-gel estimation  
 Quality coverage: 1.4x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbankdraft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html)).  
 \* NOTE: This is a "working draft" sequence. It currently  
 \* consists of 26 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1  
 4737: contig of 4737 bp in length  
 4738  
 4837: gap of unknown length  
 4838  
 4973: contig of 4636 bp in length  
 9474  
 9573: gap of unknown length  
 9474  
 9574  
 13415: contig of 3842 bp in length  
 13416  
 13515: gap of unknown length  
 13516  
 17327  
 17426: contig of 3811 bp in length  
 17327  
 23207: contig of unknown length  
 23208  
 23307: contig of 5781 bp in length  
 23308  
 27687: gap of unknown length  
 27688  
 27787: contig of 4380 bp in length  
 31391: contig of 3604 bp in length  
 31491: gap of unknown length  
 31492  
 35716: contig of 4425 bp in length  
 35717  
 35816: gap of unknown length  
 35817  
 40730: contig of 4914 bp in length  
 40731  
 40830: gap of unknown length  
 40831  
 43983: contig of 3153 bp in length  
 44083: gap of unknown length  
 47104: contig of 3021 bp in length  
 47204: gap of unknown length  
 50511: contig of 3307 bp in length  
 50512  
 50611: gap of unknown length  
 50612  
 53734: contig of 3123 bp in length  
 53735  
 53835: gap of unknown length  
 53836  
 56324: contig of 2490 bp in length  
 56325  
 56424: gap of unknown length  
 56425  
 59000: contig of 2576 bp in length  
 59001  
 59100: gap of unknown length  
 59101  
 61416: contig of 2316 bp in length  
 61417  
 64430: gap of unknown length  
 64431  
 64530: contig of 2914 bp in length  
 64531  
 64530: gap of unknown length

```

* 64531 67643: contig of 3113 bp in length
* 67644 67743: gap of unknown length
* 67744 69936: contig of 2193 bp in length
* 69937 70036: gap of unknown length
* 70037 72722: contig of 2686 bp in length
* 72723 74862: gap of unknown length
* 74863 74963: contig of 2040 bp in length
* 74963 77123: gap of unknown length
* 77124 77224: gap of unknown length
* 77224 79433: contig of 2209 bp in length
* 79433 79532: gap of unknown length
* 79533 81558: contig of 2026 bp in length
* 81559 84009: gap of unknown length
* 84010 84109: contig of 2351 bp in length
* 84110 86146: gap of unknown length
* 86146 86146: contig of 2037 bp in length.
Location/Qualifiers
1. 86146
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-402K16"

```

```

ORIGIN
Query Match          91.1%; Score 16.4; DB 2; Length 86146;
Best Local Similarity 94.4%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
    |||||
Db 10809 CGACTGTGAATCTCCAT 10792

```

```

RESULT 45
LOCUS AL731687/c 116898 bp DNA linear ROD 06-NOV-2002
DEFINITION Mouse DNA sequence from clone RP23-56120 on chromosome 11, complete
sequence.
ACCESSION AL731687
VERSION AL731687.13 GI:24816835
KEYWORDS HMG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 116898)
Tromans, A.
JOURNAL Direct Submission
Submitted (06-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 8, 2002 this sequence version replaced gi:24366489.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-56120 is from the RPCI-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

```

FEATURES
SOURCE
1. 116898
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-56120"
/clone_id="RPCI-23"
Location/Qualifiers

ORIGIN
Query Match          91.1%; Score 16.4; DB 10; Length 116898;
Best Local Similarity 94.4%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
    |||||
Db 81044 CGACTGTGAATCTCCAT 81027

```

```

RESULT 46
LOCUS AL445243 150937 bp DNA linear PRI 03-JUL-2001
DEFINITION Human DNA sequence from clone RP11-538C21 on chromosome 13 Contains
STSs and GSSs, complete sequence.
ACCESSION AL445243
VERSION AL445243.3 GI:11136908
KEYWORDS HMG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 150937)
Wall, M.
JOURNAL Direct Submission
Submitted (29-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 10, 2000 this sequence version replaced gi:11024525.

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre. Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13> IMPORTANT: This sequence is not the entire insert of clone RP11-538C21. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-538C21 is at 1 in this sequence. The true left end of clone RP11-1685 is at 150838 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an



alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-538C21 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

#### FEATURES

source  
1..150937  
/organism="Homo sapiens"  
/mol\_type="genomicDNA"  
/db\_xref="taxon:9606"  
/chromosome="13"  
/clone="RP11-538C21"  
/clone\_1fb="RP11-11.2"  
896..951  
repeat\_region  
/note="28 copies 2 mer tt 69% conserved"  
3412..3905  
/note="match: GSS: Em:AQ663918"  
3949..4172  
repeat\_region  
/note="112 copies 2 mer tt 67% conserved"  
complement(5225..5632)  
/note="match: GSS: Em:AQ012043"  
complement(5290..5643)  
/note="match: GSS: Em:AQ056175"  
complement(6845..7233)  
/note="match: GSS: Em:AQ211228"  
7242..7762  
/note="match: GSS: Em:A2519265"  
complement(8525..9078)  
/note="match: GSS: Em:AQ422480"  
complement(8631..9064)  
/note="match: STS: Em:G53065  
match: GSS: Em:B76751"  
10605..10640  
repeat\_region  
/note="18 copies 2 mer aa 88% conserved"  
19650..19717  
repeat\_region  
/note="34 copies 2 mer tc 85% conserved"  
complement(28596..28945)  
/note="match: GSS: Em:AQ458456"  
29096..29557  
/note="match: GSS: Em:AQ772130"  
complement(30788..31290)  
/note="match: STS: Em:G57193  
match: GSS: Em:AQ341794"  
complement(34113..34607)  
/note="match: GSS: Em:AQ194187"  
complement(34157..34604)  
/note="match: GSS: Em:AQ814840"  
complement(34168..34607)  
/note="match: GSS: Em:AQ198600"  
34713..34802  
repeat\_region  
/note="WT2FA repeat: matches 81..181 of consensus"  
38474..38849  
repeat\_region  
/note="THE1C repeat: matches 1..371 of consensus"  
41672..42240  
repeat\_region  
/note="THE1B-INTERNAL repeat: matches 1..615 of consensus"  
44388..44427  
repeat\_region  
/note="20 copies 2 mer ac 95% conserved"  
complement(44392..44726)  
/note="match: GSS: Em:AQ488076"  
44846..45259  
/note="match: GSS: Em:B90508"  
complement(45859..46009)  
/note="match: STS: Em:AF003650"  
46159..46418  
repeat\_region  
/note="MER52A repeat: matches 528..750 of consensus"  
46440..46617  
repeat\_region  
/note="LTR28 repeat: matches 1..178 of consensus"  
47959..48082  
/note="match: STS: Em:G27256"

repeat\_region  
48084..48147  
/note="32 copies 2 mer ta 70% conserved"  
misc\_feature  
48088..48301  
/note="match: STS: Em:G27256"  
misc\_feature  
49323..49712  
/note="match: GSS: Em:AQ375299"  
53872..53931  
repeat\_region  
/note="MER67D repeat: matches 445..506 of consensus"  
54086..54442  
repeat\_region  
/note="MER67C repeat: matches 1..367 of consensus"  
54824..55503  
misc\_feature  
/note="match: GSS: Em:AQ055122"  
57402..57435  
repeat\_region  
/note="17 copies 2 mer ct 91% conserved"  
63756..64537  
/note="match: GSS: Em:AQ537963"  
complement(63874..64285)  
/note="match: GSS: Em:AQ465707"  
66074..66115  
repeat\_region  
/note="21 copies 2 mer tt 85% conserved"  
67279..67780  
/note="match: GSS: Em:AQ765920"  
complement(68370..68622)  
/note="match: GSS: Em:A2391162"  
complement(68391..69107)  
/note="match: GSS: Em:AQ527862"  
complement(68437..68626)  
/note="match: GSS: Em:AQ807981"  
complement(68514..68639)  
/note="match: GSS: Em:AQ456087"  
69124..69559  
misc\_feature  
/note="match: GSS: Em:AQ881502"  
71416..71505  
repeat\_region  
/note="45 copies 2 mer ta 73% conserved"  
71511..71871  
repeat\_region  
/note="L1P3 repeat: matches 5783..6146 of consensus"  
78843..79653  
repeat\_region  
/note="T1ger3b repeat: matches 407..1231 of consensus"  
79964..80405  
repeat\_region  
/note="T1ger3b repeat: matches 1..407 of consensus"  
82762..83871  
repeat\_region  
/note="L1PA15 repeat: matches 604..1684 of consensus"  
83872..84741  
repeat\_region  
/note="MER1D repeat: matches 3..897 of consensus"  
84742..86381  
repeat\_region  
/note="L1PA15 repeat: matches 1684..3280 of consensus"  
86682..89269  
repeat\_region  
/note="L1PA15 repeat: matches 3280..6157 of consensus"  
91758..91799  
repeat\_region  
/note="21 copies 2 mer ac 95% conserved"  
91815..91870  
repeat\_region  
/note="28 copies 2 mer ac 78% conserved"  
92029..92491  
repeat\_region  
/note="LTR33 repeat: matches 38..521 of consensus"  
95335..96172  
repeat\_region  
/note="MER76 repeat: matches 1..666 of consensus"  
96190..96217  
repeat\_region  
/note="14 copies 2 mer cc 89% conserved"  
96233..96254  
repeat\_region  
/note="MER76 repeat: matches 667..688 of consensus"  
97703..98137  
/note="match: GSS: Em:AQ077541"  
98891..99319  
repeat\_region  
/note="L1P repeat: matches 4766..5195 of consensus"  
102030..102067  
repeat\_region  
/note="19 copies 2 mer tt 78% conserved"  
109590..109767  
repeat\_region  
/note="89 copies 2 mer tt 66% conserved"  
109797..110092  
repeat\_region  
/note="MSTC repeat: matches 1..289 of consensus"  
113217..113786  
misc\_feature  
/note="match: GSS: Em:AQ017010"  
complement(114392..114852)

```

misc_feature      /note="match: GSS: Em:B94847"
                  /note="match: GSS: Em:AQ119585"
misc_feature      /note="match: GSS: Em:AQ081553"
misc_feature      /note="match: GSS: Em:AQ33498 Em:AQ431610"
misc_feature      /note="match: GSS: Em:AZ497064"
misc_feature      /note="match: GSS: Em:AQ319097"
misc_feature      /note="match: GSS: Em:G10072"
repeat_region     /note="54 copies 2 mer ta 644 conserved"

Query Match      91.1%; Score 16.4; DB 9; Length 150937;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 CGACTGTGAATCCTCCAT 18

Db 41556 CGACTGTGAATCCTCCAT 41573

```

RESULT 47
AC102741/c      163447 bp      DNA      linear      HTG 20-MAR-2004
LOCUS           Mus musculus chromosome 17 clone RP24-343110 map 17, *** SEQUENCING
DEFINITION     AC102741
ACCESSION      AC102741
VERSION        GI:45597679
KEYWORDS       HTG; HTGS_PHASE1; HTGS_FULUTOP; HTGS_ACTIVEFIN.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 163447)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 17, clone RP24-343110
Unpublished
2 (bases 1 to 163447)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Baetsen,V., Boguslavsky,L., Bouknight,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Govette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,
Lamaze,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,W., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Stratus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

TITLE Direct Submission  
JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
REFERENCE 3 (bases 1 to 163447)  
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Boguslavsky,L., Bouknight,B., Camarata,J., Chang,J., Chapel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,

TITLE  
JOURNAL  
COMMENT

Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Haefl,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Liu,A., Mabbitt,R., Maclean,C., MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (20-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 20, 2004 this sequence version replaced gi:28630066.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center Project name: L19396  
Center Clone name: 343\_I10

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 16709: contig of 16709 bp in length
* 16710 16809: gap of 100 bp
* 16810 20326: contig of 3517 bp in length
* 20327 20426: gap of 100 bp
* 20427 26699: contig of 6273 bp in length
* 26700 26700: gap of 100 bp
* 26800 51907: contig of 25108 bp in length
* 51908 52007: gap of 100 bp
* 52008 97610: contig of 45603 bp in length
* 97611 97710: gap of 100 bp
* 97711 134460: contig of 36750 bp in length
* 134461 134560: gap of 100 bp
* 134561 150090: contig of 15330 bp in length
* 150091 150190: gap of 100 bp
* 150191 163447: contig of 13257 bp in length.
Location/Qualifiers
1. 163447
/organism="Mus musculus"
/mol_type="Genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP24-343110"
/clone_lib="RPCI-24 Male Mouse BAC"

```

ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 163447;  
Best Local Similarity 94.4%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGACTGTGAATCCTCCAT 18  
Db 106744 CGACTGTGAATCCTCCAT 106727

RESULT 48  
BX324147 172568 bp DNA linear HTG 24-MAY-2003  
LOCUS  
DEFINITION Danio rerio clone DKEX-217G15, WORKING DRAFT SEQUENCE.  
ACCESSION BX324147  
VERSION BX324147.2 GI:31075194  
KEYWORDS HTG, HTGS\_PHASE2, HTGS\_DRAFT, HTGS\_FULFILL.  
SOURCE Danio rerio (zebrafish)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 172568)  
McLaren, S.  
Direct Submission  
Submitted (23-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On May 25, 2003 this sequence version replaced gi:30024560.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: zfish-help@sanger.ac.uk  
----- Project Information  
Center project name: zK217G15  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 172418 bases at least Q40  
Consensus quality: 172539 bases at least Q30  
Consensus quality: 172564 bases at least Q20  
Insert size: 172568; sum-of-contigs  
Insert size: 162997; 3.0% error; agarose-fp  
Quality coverage: 5.65x in Q20 bases; sum-of-contigs Quality  
coverage: 6.03x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 172568: contig of 172568 bp in length.  
Location/Qualifiers  
1. 172568  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEX-217G15"  
/clone\_lib="DanioKey"  
1. 172568  
/note="assembly\_fragment:01051"  
misc\_feature  
ORIGIN  
Query Match 91.1%; Score 16.4; DB 2; Length 172568;  
Best Local Similarity 94.4%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 CGACTGTAATCTTCAT 18  
DB 106266 CGACTGTAATCTTCGT 106283  
RESULT 49  
AC074146 201602 bp DNA linear HTG 15-JUL-2000  
LOCUS  
DEFINITION Mus musculus clone RP23-107G3, WORKING DRAFT SEQUENCE, 38 unordered  
pieces.

ACCESSION AC074146  
VERSION AC074146.1 GI:9211256  
KEYWORDS HTG, HTGS\_PHASE1, HTGS\_DRAFT.  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 201602)  
DOE Joint Genome Institute.  
Sequencing of Mouse  
Unpublished  
2 (bases 1 to 201602)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
----- Project Information  
Center Project Name: 0  
Center clone name: RP23-107G3  
-----  
Summary Statistics  
Consensus quality: 134464 bases at least Q40  
Consensus quality: 159739 bases at least Q30  
Consensus quality: 170106 bases at least Q20  
Estimated insert size: 181300; agarose-fp estimation  
Estimated insert size: 197902; sum-of-contigs estimation  
Quality coverage: 3.41 in Q20 bases; agarose-fp estimation  
Quality coverage: 3.12 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 38 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1675: contig of 1675 bp in length  
\* 1676 1775: gap of unknown length  
\* 1776 3291: contig of 1516 bp in length  
\* 3292 3391: gap of unknown length  
\* 3392 4432: contig of 1041 bp in length  
\* 4433 4532: gap of unknown length  
\* 4533 6053: contig of 1521 bp in length  
\* 6054 6153: gap of unknown length  
\* 6154 7161: contig of 1008 bp in length  
\* 7162 7261: gap of unknown length  
\* 7262 8417: contig of 1156 bp in length  
\* 8418 8517: gap of unknown length  
\* 8518 9800: contig of 1283 bp in length  
\* 9801 9900: gap of unknown length  
\* 9901 10926: contig of 1026 bp in length  
\* 10927 11026: gap of unknown length  
\* 11027 12062: contig of 1036 bp in length  
\* 12063 12162: gap of unknown length  
\* 12163 14110: contig of 1948 bp in length  
\* 14111 14210: gap of unknown length  
\* 14212 14211: gap of unknown length  
\* 14211 15534: contig of 1324 bp in length  
\* 15535 15634: gap of unknown length  
\* 15635 17316: contig of 1682 bp in length  
\* 17317 17416: gap of unknown length  
\* 17417 20982: contig of 3566 bp in length  
\* 20983 21083: gap of unknown length  
\* 21083 24936: contig of 3854 bp in length  
\* 24937 25036: gap of unknown length  
\* 25037 28124: contig of 3088 bp in length  
\* 28125 28224: gap of unknown length  
\* 28225 30948: contig of 2724 bp in length  
\* 30949 31048: gap of unknown length

```

* 31049 34011: contig of 2963 bp in length
* 34012 34111: gap of unknown length
* 34112 37382: contig of 3271 bp in length
* 37383 37482: gap of unknown length
* 37483 41136: contig of 3654 bp in length
* 41137 41236: gap of unknown length
* 41237 45754: contig of 4518 bp in length
* 45755 45854: gap of unknown length
* 45855 48474: contig of 2620 bp in length
* 48475 48574: gap of unknown length
* 48575 52059: contig of 3485 bp in length
* 52060 52159: gap of unknown length
* 52160 55993: contig of 3834 bp in length
* 55994 56093: gap of unknown length
* 56094 59136: contig of 3043 bp in length
* 59137 59236: gap of unknown length
* 59237 63015: contig of 3779 bp in length
* 63016 63115: gap of unknown length
* 63116 66917: contig of 3802 bp in length
* 66918 67017: gap of unknown length
* 67018 72225: contig of 5208 bp in length
* 72226 72325: gap of unknown length
* 72326 80793: contig of 8468 bp in length
* 80794 80893: gap of unknown length
* 80894 88129: contig of 7236 bp in length
* 88130 88229: gap of unknown length
* 88230 95766: contig of 7537 bp in length
* 95767 95866: gap of unknown length
* 95867 104692: contig of 8826 bp in length
* 104693 104792: gap of unknown length
* 104793 113453: contig of 8661 bp in length
* 113454 113553: gap of unknown length
* 113554 121609: contig of 8056 bp in length
* 121610 121709: gap of unknown length
* 121710 132117: contig of 10408 bp in length
* 132118 132217: gap of unknown length
* 132218 143836: contig of 11619 bp in length
* 143837 143936: gap of unknown length
* 143937 153823: contig of 9887 bp in length
* 153824 153923: gap of unknown length
* 153924 174899: contig of 20976 bp in length
* 174900 174999: gap of unknown length
* 175000 201602: contig of 26603 bp in length.

FEATURES
Source
location/Qualifiers
1..201602
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-107G3"
/clone_lib="RP23 mouse BAC library 23"

ORIGIN
Query Match 91.1%; Score 16.4; DB 2; Length 201602;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTCATCTCCAT 18
Db 197993 CGACTGTCATCTCCAT 197976

RESULT 50
AC148732/c
LOCUS AC148732 214262 bp DNA linear HTG 31-MAR-2004
DEFINITION Callithrix jacchus clone CH259-164P15, WORKING DRAFT SEQUENCE, 11
unordered pieces.
AC148732
AC148732.1 GI:45860953
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Callithrix jacchus (white-tufted-ear marmoset)
SOURCE Callithrix jacchus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;

```

```

REFERENCE
AUTHORS
1 (bases 1 to 214262)
Antonelli,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Grantie,S., Guan,X., Gupta,D.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Idol,J.R., Jones,C., Karlins,B., Kim,H., Kwong,P., Lartie,P.,
Larson,S., Lee-Jin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Paguirigan,C., Pagnier,M.E., Prasad,A., Puri,O.,
Redix-Duque,N., Schandler,K., Schueler,M.G., Shah,K., Simon,C.,
Stentropop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.D.,
Wehenby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 214262)
Green,E.D.
Direct Submission
Submitted (31-MAR-2004) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Galthersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoehngri.nih.gov
----- Project Information
Center project name: 164P15
Center clone name: 164P15
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 209757 bases at least Q40
Consensus quality: 210672 bases at least Q30
Consensus quality: 211371 bases at least Q20
Insert size: 213000; agarose-fp
Insert coverage: 8.19x in Q20 bases; sum-of-contigs
Quality coverage: 8.18x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2648: contig of 2648 bp in length
2649 2749 2748: gap of unknown length
2749 4920: contig of 2172 bp in length
4921 5020: gap of unknown length
5021 7682: contig of 2662 bp in length
7683 7782: gap of unknown length
7783 22178: contig of 14396 bp in length
22179 22278: gap of unknown length
22279 29583: contig of 7305 bp in length
29584 29683: gap of unknown length
29684 40166: contig of 10483 bp in length
40167 40266: gap of unknown length
40267 57686: contig of 17420 bp in length
57687 57787: gap of unknown length
57788 73989: contig of 16003 bp in length
73990 74089: gap of unknown length
74090 97609: contig of 23520 bp in length
97610 97709: gap of unknown length
97710 144022: contig of 46313 bp in length
144023 214262: gap of unknown length
144123 location/Qualifiers
1..214262
/organism="Callithrix jacchus"
/mol_type="genomic DNA"

```

```
/db_xref="taxon:9483"
/clone="CH259-164P15"
/clone_lib="CH259"
/note="BAC resource: http://bacpac.chori.org/"
misc_feature
1..2648
/note="assembly_fragment"
misc_feature
2749..4920
/note="assembly_fragment"
misc_feature
5021..7682
/note="assembly_fragment"
misc_feature
7783..22178
/note="assembly_fragment"
misc_feature
22279..29583
/note="assembly_fragment"
misc_feature
clone_end:SP6
vector_side:right"
29684..40166
/note="assembly_fragment"
misc_feature
40267..57686
/note="assembly_fragment"
misc_feature
57787..73989
/note="assembly_fragment"
misc_feature
74090..97609
/note="assembly_fragment"
misc_feature
97710..144022
/note="assembly_fragment"
clone_end:r7
vector_side:left"
144123..214262
/note="assembly_fragment"
```

## ORIGIN

```
Query Match          91.1%; Score 16.4; DB 2; Length 214262;
Best Local Similarity 94.4%; Pred.No.1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CGACTGTGATCTCCAT 18
        |||||
Db      2777 CGACTGTGATCCCCAT 2760
```

Search completed: January 14, 2005, 16:03:08  
Job time : 984.263 secs

**THIS PAGE BLANK (USPTO)**